



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number 116405**

**TO: James Schultz**  
**Location: REM/2D182C18**  
**Art Unit: 1635**  
**Monday, March 01, 2004**

**Case Serial Number: 09/904568**

**From: Toby Port**  
**Location: Biotech-Chem Library**  
**Remsen 1A59**  
**Phone: 571-272-2523**

**toby.port@uspto.gov**

### **Search Notes**

Dear Examiner Schultz,

Here are the results of your search.  
Please feel free to contact me if you have any questions.

Toby Port



# STIC SEARCH RESULTS FEEDBACK FORM

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor  
571-272-2507 Remsen E01 D86

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library Remsen Bldg.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 29, 2004, 09:53:57 ; Search time 303.726 Seconds  
(without alignments)  
9413.223 Million cell updates/sec

Title: US-09-904-568-3\_COPY\_294\_966

Perfect score: 673  
Sequence: 1 aatgctcgtctgtg9399ctg.....gtgagcgagactgcagagac 673

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapexr 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 3185356

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	3.4	48	2	AAT89912 Human ubi
2	22.4	3.3	36	3	AAA35922 Permuten
3	21.4	3.2	43	2	AAV31116 H chain v
4	21.4	3.2	43	3	AAZ44243 Murine CD
5	21.4	3.2	50	4	AAI28467 Human SNP
6	21.2	3.2	44	2	AAI19291 Human gta
7	21.2	3.2	50	4	AAI33664 Human SNP
8	20.8	3.1	50	4	AAI33563 Human SNP
9	20.6	3.1	40	6	ABZ49057 Human ALD
10	20.6	3.1	40	6	ABZ46527 Human ALD
11	20.6	3.1	41	2	AAI19312 Human gta
12	20.2	3.0	42	4	AAI68438 Primer_17
13	20	3.0	41	6	ABZ46702 Human N-m
14	20	3.0	41	6	ABZ43362 Human N-m
15	20	3.0	45	7	ABT34022 Human pig
16	20	3.0	50	6	ABZ03550 Human leu
17	19.8	2.9	45	3	AAQ94175 Transmem
18	19.8	2.9	45	3	AAQ05785 Streptavi
19	19.8	2.9	48	3	AAI43970 C allele
20	19.8	2.9	49	7	ACD18524 Human zin
21	19.8	2.9	50	4	AAI30588 Human SNP
22	19.8	2.9	50	4	AAI28730 Human SNP
23	19.6	2.9	41	6	ABZ45511 Human ATP

C	24	19.6	2.9	41	6	ABZ46917 Human ATP
C	25	19.6	2.9	45	3	AAI70390 Site dire
C	26	19.6	2.9	46	2	AAQ68290 Diabody M
C	27	19.6	2.9	48	4	ABK07023 Human NOG
C	28	19.6	2.9	50	6	AAI28862 Human SNP
C	29	19.6	2.9	50	6	ABZ04150 Human leu
C	30	19.4	2.9	31	2	AAI60199 PCR prime
C	31	19.4	2.9	48	2	AAV56403 Human ICA
C	32	19.4	2.9	48	2	AAV69175 Chimeric
C	33	19.4	2.9	48	2	AAI21873 Primer fo
C	34	19.4	2.9	48	3	AAZ24317 Humanised
C	35	19.4	2.9	48	3	AAI97145 PCR prime
C	36	19.4	2.9	48	3	AAI08291 Chimeric
C	37	19.4	2.9	48	3	AAI49711 C allele
C	38	19.4	2.9	48	6	ABK09334 Humanised
C	39	19.4	2.9	50	2	AAV68378 Clone #4
C	40	19.4	2.9	50	4	AAI73981 Human g11
C	41	19.4	2.9	50	6	ABQ72370 PCR prime
C	42	19.4	2.9	50	7	ACC00329 Oligonuc1
C	43	19.2	2.9	36	3	AAI35507 Permuten
C	44	19.2	2.9	47	7	ACD28216 Spirochet
C	45	19.2	2.9	50	4	AAI33865 Human SNP

## ALIGNMENTS

RESULT 1  
AAT89912/c  
ID AAT89912 standard; DNA; 48 BP.  
XX  
AC AAT89912;  
XX  
DT 14-APR-1998 (first entry)  
XX  
DE Human ubiquitin carrier polypeptide UbCH10 primer HSEAUC.  
XX  
KM Ubiquitin carrier protein; UbCH10; human; cyclin A; cyclin B;  
KM Ubiquitination; cell cycle; mitosis; PCR; primer; AUI epitope;  
KM dominant negative mutant; ss.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
PN MO9737027-A1.  
XX  
PD 09-OCT-1997.  
XX  
PF 31-MAR-1997; 97WO-US005296.  
XX  
PR 01-APR-1996; 96US-0014492P.  
PR 18-MAR-1997; 97US-00820639.  
PR 31-MAR-1997; 97US-00828533.  
XX  
PA (HARD) HARVARD COLLEGE.  
XX  
XX Ruderman JV, Hershko A, Kirschner MW, Townsley F, Aristarkov A;  
XX Eytan E, Yu H;  
XX WPI; 1997-503112/46.  
PT New ubiquitin carrier polypeptide(s) - are involved in ubiquitination of  
PT cyclin A and cyclin B and are useful to screen for new inhibitors of cell  
PT proliferation.  
XX  
XX Example 8; Page 86; 138pp; English.  
CC PCR antisense primer HSEAUC (AAT89912) encodes the last 6 amino acids of  
CC the open reading frame (see AAT89912) of UbCH10 (see AAI31277), a novel  
CC human ubiquitin carrier protein that is involved in the ubiquitination  
CC and degradation of mitotic cyclins. HSEAUC also encodes the amino acids  
CC DTYR1 (AUI epitope) and includes 2 stop codons and an HindIII  
CC restriction site. It was used with sense primer HSEN (see AAT89910) to

CC add the AVI epitope to the C-terminus of wild-type UbcH10 and its  
CC Cys114Ser mutant (see AAW31279). Tagged proteins were expressed in E.  
CC coli and used in tests to examine the in vivo and in vitro activity of  
CC UbcH10 and its dominant negative mutant  
XX  
SQ Sequence 48 BP; 9 A; 9 C; 17 G; 13 T; 0 U; 0 Other;  
Query Match 3.4%; Score 23; DB 2; Length 48;  
Best Local Similarity 83.9%; Pred. No. 5.1e+04;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 304 CACGACCTGAAAGCTGACACCTTCAGGAC 334  
DB 47 CACGACGAGAGCCGACACCTACAGGTAC 17  
RESULT 2  
ID AAW35922/c  
AAW35922 standard; DNA; 36 BP.  
XX  
XX AAA35922;  
AC  
XX  
XX 26-JUL-2000 (first entry)  
DT  
XX  
XX Permutin linker encoding nucleotide sequence RGS152.  
DE  
XX  
XX Biologically-activated circularly-permuted protein; permutin; linker;  
KW permutin library generation; therapeutic property; antigen;  
KW immunotherapy; improve bio-distribution; half life; ss.  
XX  
XX OS Synthetic.  
XX  
XX WO200018905-A1.  
XX  
XX PD 06-APR-2000.  
XX  
XX PF 24-SEP-1999; 99WO-US020891.  
XX  
XX PR 25-SEP-1998; 98US-0101908P.  
XX  
XX PA (SEAR) SEARLE & CO G. D.  
XX  
XX PI Lee SC;  
XX  
XX DR WPI; 2000-293145/25.  
XX  
XX PT Preparation of biologically-activated circularly-permuted proteins by  
PT scanning permutagenesis for generating libraries of permutins with  
PT improved therapeutic properties.  
XX  
XX PS Claim 11; Page 44; 100pp; English.  
XX  
CC The preparation of biologically-activated circularly-permuted proteins  
CC (permutins) comprises the use of a method comprising making a series of  
CC circularly permuted genes. The circularly permuted genes are inserted  
CC into a display vector, where they are expressed so that the proteins they  
CC encode are presented on the surface of the display vector. A library of  
CC display vectors presenting the expressed circularly permuted proteins is  
CC generated. A target protein that can bind a biologically active  
CC circularly permuted protein can be used to affinity select the presenting  
CC display vectors. The selected display vectors can be isolated and  
CC analyzed to identify the presented circularly-permuted protein. The  
CC permutins conform to the structure of a parent protein consisting of a  
CC segment derived from the carboxy portion of the parent protein, a segment  
CC derived from the amino terminus of the parent protein, and a linker or  
CC chemical bond linking the amino and carboxy terminal derived portions.  
CC Nucleotide sequences AAW3576-A35943 encode linkers used to create the  
CC permutins of the invention. The method is used to generate libraries of  
CC permutins with improved therapeutic properties compared to their parent  
CC molecules. Permutins with little or no activity may be used as antigens  
CC for producing antibodies which are used in immunology or immunotherapy as  
CC probes or intermediates used to construct other useful permutins.  
CC Permutins have improved biological and therapeutic properties compared

CC to their two individual components due to alterations in bio-distribution  
CC or half-life  
XX  
SQ Sequence 36 BP; 5 A; 15 C; 10 G; 6 T; 0 U; 0 Other;  
Query Match 3.3%; Score 22.4; DB 3; Length 36;  
Best Local Similarity 81.2%; Pred. No. 6.6e+04;  
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 425 GCCGAGCAGCAGGAGGAGCTGCTGCCGATGT 456  
DB 35 GCCGAGCAGCAGGAGGAGCTGCTGCCGATAT 4  
RESULT 3  
ID AAV31116  
AAV31116 standard; DNA; 43 BP.  
XX  
XX AAV31116;  
AC  
XX  
XX 02-SEP-1998 (first entry)  
DT  
XX  
XX H chain variable region from mouse anti-MY-10 PCR primer 1.  
DE  
XX  
XX Mouse; anti-MY-10; H chain variable region; L chain variable region;  
KW antibody; heavy; light; CDR; human CD34 antigen; hybridoma; anti-CD34;  
KW PCR primer; ss.  
XX  
XX OS Synthetic.  
XX  
XX Mus sp.  
XX  
XX JP10155489-A.  
XX  
XX PD 16-JUN-1998.  
XX  
XX PF 27-NOV-1996; 96JP-00331647.  
XX  
XX PR 27-NOV-1996; 96JP-00331647.  
XX  
XX PA (ASAH) ASAH KASEI KOGYO KK.  
XX  
XX PA (ASAH) ASAH MEDICAL CO LTD.  
XX  
XX DR WPI; 1998-391043/34.  
XX  
XX PT Recombinant antibody against human CD34 - and nucleic acid encoding it,  
PT used for efficient production of the antibody.  
XX  
XX PS Example 4; Page 7; 15pp; Japanese.  
XX  
CC The present invention describes a nucleic acid which encodes an anti-MY-  
CC 10 antibody. The present sequence represents a PCR primer for the H chain  
CC variable region of the antibody. Also described is a method for the  
CC production of a recombinant antibody in which the above nucleic acid is  
CC used to produce an antibody which combines with human CD34 antigen by  
CC gene recombination. The anti-CD34 antibody gene is used to produce  
CC recombinant anti-CD34 antibodies efficiently, which can be used in  
CC pharmaceuticals  
XX  
SQ Sequence 43 BP; 8 A; 15 C; 16 G; 4 T; 0 U; 0 Other;  
Query Match 3.2%; Score 21.4; DB 2; Length 43;  
Best Local Similarity 71.8%; Pred. No. 1.2e+05;  
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
QY 513 CCTGACGACGAGCTCTGATGACGCTGAAGCTTTAG 551  
DB 5 CCTGACGACGAGCTCTGATGACGCTGAAGCTGAG 43  
RESULT 4  
ID AAZ44243  
AAZ44243 standard; DNA; 43 BP.  
XX

AC AA244243;  
 XX  
 DT 31-MAR-2000 (first entry)  
 XX  
 DE Murine CD4/CD34 recognizing antibody primer 7.  
 XX  
 KW Cluster differentiation; cell separation; antibody; CD4; CD34; leukemia;  
 KW hematopoietic; undifferentiated; lymphocyte; bone marrow transplantation;  
 KW HIV infection; autoimmune disease; murine; primer; ss.  
 XX  
 OS Mus sp.  
 XX  
 PN WO9961629-A1.  
 XX  
 PD 02-DEC-1999.  
 XX  
 PF 24-MAY-1999; 99WO-JP002711.  
 XX  
 PR 25-MAY-1998; 98JP-00159957.  
 PR 26-MAY-1998; 98JP-00163023.  
 XX  
 PA (ASAH ) ASAH KASEI KOGYO KK.  
 PA (ASAH ) ASAH MEDICAL CO LTD.  
 XX  
 PI Ono M, Soka T, Morimoto I, Miyamura K;  
 XX  
 DR WPI, 2000-086720/07.  
 XX  
 PT Devices containing antibodies recognising CD4 or CD34 and their use for  
 PT the separation of CD4 or CD34 positive cells.  
 XX  
 PS Example 21; Page 54; 11pp; Japanese.  
 XX  
 CC This invention describes a novel device (I) for separating cluster  
 CC differentiation (CD)-positive cells using a recombinant (chimeric or  
 CC single-chain) antibody recognising CD4 or CD34. The devices are useful  
 CC for the separation of CD4 or CD34 positive cells, which is useful for the  
 CC collection of hematopoietic undifferentiated cells, elimination of  
 CC lymphocytes from cells to be used in bone marrow transplantation, the  
 CC detection of leukemic cells and the production of medicinal compositions  
 CC for the treatment of HIV infection and autoimmune diseases. AA244237-  
 CC 244246 represent primers used to illustrate the method of the invention  
 CC  
 SO Sequence 43 BP; 8 A; 15 C; 16 G; 4 T; 0 U; 0 Other;  
 XX  
 QY Query Match 3.2%; Score 21.4; DB 3; Length 43;  
 Best Local Similarity 71.8%; Pred. No. 1.2e+05;  
 Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
 XX  
 Db 513 CCCTGCAGCCGAGGCTCTGATGACGCTGAAGCTTTTCAG 551  
 5 CCCAGCCGCGCCATGCGCCAGGTGACGTGAAGCAGTCAAG 43  
 XX  
 RESULT 5  
 AA128467  
 ID AA128467 standard; DNA; 50 BP.  
 XX  
 AC AA128467;  
 XX  
 DT 24-JAN-2002 (first entry)  
 XX  
 DE Human SNP oligonucleotide #1675.  
 XX  
 KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;  
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
 KW multifactorial disease; autoimmune disease; infection;  
 KW nervous system disease; ss.  
 XX

OS Homo sapiens.  
 XX  
 PN WO200147944-A2.  
 XX  
 PD 05-JUL-2001.  
 XX  
 PF 28-DEC-2000; 2000WO-US035498.  
 XX  
 PR 28-DEC-1999; 99US-0173419P.  
 PR 27-DEC-2000; 2000US-00173419.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shimkets RA, Leach M;  
 XX  
 DR WPI, 2001-465210/50.  
 XX  
 PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
 PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,  
 PT autoimmune diseases and infections.  
 XX  
 PS Claim 1; Page 1859; 4143pp; English.  
 XX  
 CC The present invention relates to oligonucleotides encoding polymorphic  
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
 CC histones, kinases, colony stimulating factors, complement related  
 CC proteins, cytochromes, kinesin, cytokines, interferons, interleukins, G-  
 CC protein coupled receptors and thioesterases. The present sequence is one  
 CC such oligonucleotide. The oligonucleotides and the peptides encoded by  
 CC them may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate expression of the proteins listed above.  
 CC Disorders that may be prevented, diagnosed and/or treated include  
 CC multifactorial diseases with a genetic component, such as autoimmune  
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,  
 CC leukemia), diseases of the nervous system and an infection of pathogenic  
 CC organisms  
 CC  
 SO Sequence 50 BP; 9 A; 14 C; 22 G; 5 T; 0 U; 0 Other;  
 XX  
 QY Query Match 3.2%; Score 21.4; DB 4; Length 50;  
 Best Local Similarity 66.0%; Pred. No. 1.2e+05;  
 Matches 31; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
 XX  
 Db 56 AGTGGCCGAGTGAGTTCTTGGGGCCCGAGCTTCTCCAGAGGTGAC 102  
 2 AGGGGAGAGCTGGGAGGTGGGAGCCCACTCTGCCAGAGGTGGC 48  
 XX  
 RESULT 6  
 AA119291/c  
 ID AA119291 standard; DNA; 44 BP.  
 XX  
 AC AA119291;  
 XX  
 DT 17-MAY-1999 (first entry)  
 XX  
 DE Human granulocyte colony-stimulating factor antisense primer oligo 1.  
 XX  
 KW Human; granulocyte colony-stimulating factor; G-CSF; hg-CSF;  
 KW haematopoiesis disorder; primer; ss.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9853072-A1.  
 XX  
 PD 26-NOV-1998.  
 XX  
 PF 21-MAY-1998; 98WO-KR000125.  
 XX

PR 22-MAY-1997; 97KR-00020054.  
XX (KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.  
PA (JEIL-) JEIL PHARM CO LTD.  
XX  
PI Jun H, Choi B, Lee Y, Sohn M;  
XX WPI, 1999-131691/11.  
DR  
PT New human granulocyte colony-stimulating factor (hg-CSF) containing a  
PT peptide at the N-terminus - useful in the treatment of haematopoiesis  
PT disorders.  
XX  
PS Example 1; Fig 1; 19pp; English.  
XX  
CC The present invention describes the N-terminal peptide of human  
CC granulocyte colony-stimulating factor (hg-CSF). The present invention  
CC also describes: (1) recombinant plasmid pYHM-G-CSF containing the cDNA  
CC for the peptide; and (2) E. coli BL21 (pYHM-G-CSF) (KCTC 0477BP)  
CC transformed with the plasmid. The plasmid and E. coli strain are used to  
CC produce high yields of highly purified hg-CSF, which when refolded has  
CC comparable biological activity of G-CSF. hg-CSF is used to treat  
CC haematopoiesis disorders. The method of producing hg-CSF is less time-  
CC consuming and, therefore, more economical than previous methods because  
CC the protein is expressed in high yields without the need for further  
CC processing steps. AAX19290 to AAX19315 represent primer used in the  
CC synthesis of hg-CSF from an example of the present invention  
SQ Sequence 44 BP; 9 A; 10 C; 20 G; 5 T; 0 U; 0 Other;  
XX  
Query Match 3.2%; Score 21.2; DB 2; Length 44;  
Best Local Similarity 69.0%; Pred. No. 1.3e+05;  
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
QY 493 CACCAGTCCCTGCGCTGCTCCCTGAGCGCGAGCTCCGAT 534  
DB 43 CTCACGCGGTCTCTCCAGCTCCCTGCGCGAGAGCTTCGCT 2  
RESULT 7  
ID AAL33664/c  
XX AAL33664 standard; DNA; 50 BP.  
XX  
AC AAL33664;  
XX  
DT 24-JAN-2002 (first entry)  
XX  
DE Human SNP oligonucleotide #6872.  
XX  
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
KW neuroprotective; antimicrobial; gene therapy; vaccine; amyase; cancer;  
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
KW multifactorial disease; autoimmune disease; infection;  
KW nervous system disease; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200147944-A2.  
XX  
PD 05-JUL-2001.  
XX  
PF 28-DEC-2000; 2000MO-US035498.  
XX  
PR 28-DEC-1999; 99US-0173419P.  
XX 27-DEC-2000; 2000US-00173419.  
XX  
PA (CURA-) CURAGEN CORP.  
XX Shimkets RA, Leach M;  
PI  
XX

DR WPI; 2001-465210/50.  
XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,  
PT autoimmune diseases and infections.  
XX  
PS Claim 1; Page 3347; 4143pp; English.  
XX  
CC The present invention relates to oligonucleotides encoding polymorphic  
CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
CC histones, kinases, colony stimulating factors, complement related  
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins, G-  
CC protein coupled receptors and thioesterases. The present sequence is one  
CC such oligonucleotide. The oligonucleotides and the peptides encoded by  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate expression of the proteins listed above.  
CC Disorders that may be prevented, diagnosed and/or treated include  
CC multifactorial diseases with a genetic component, such as autoimmune  
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,  
CC leukaemia), diseases of the nervous system and an infection of pathogenic  
CC organisms  
XX  
SQ Sequence 50 BP; 12 A; 15 C; 13 G; 10 T; 0 U; 0 Other;  
XX  
Query Match 3.2%; Score 21.2; DB 4; Length 50;  
Best Local Similarity 69.0%; Pred. No. 1.4e+05;  
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
QY 76 GGGGGCCCAAGCTTCTCCAGAGGTGGCAGCATGCGCCGCT 117  
DB 46 GGTGACGCTTCTCTCCAGACTTGGCAGCAGACGACCGTT 5  
RESULT 8  
ID AAL33563  
XX AAL33563 standard; DNA; 50 BP.  
XX  
AC AAL33563;  
XX  
DT 24-JAN-2002 (first entry)  
XX  
DE Human SNP oligonucleotide #6771.  
XX  
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
KW neuroprotective; antimicrobial; gene therapy; vaccine; amyase; cancer;  
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
KW multifactorial disease; autoimmune disease; infection;  
KW nervous system disease; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200147944-A2.  
XX  
PD 05-JUL-2001.  
XX  
PF 28-DEC-2000; 2000MO-US035498.  
XX  
PR 28-DEC-1999; 99US-0173419P.  
XX 27-DEC-2000; 2000US-00173419.  
XX  
PA (CURA-) CURAGEN CORP.  
XX Shimkets RA, Leach M;  
PI  
XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
PT

PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,  
XX autoimmune diseases and infections.  
PS Claim 1, Page 33:5, 4143pp; English.  
XX  
CC The present invention relates to oligonucleotides encoding polymorphic  
CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
CC histones, kinases, colony stimulating factors, complement related  
CC proteins, cytochromes, kinases, cytokines, interleukins, G-  
CC protein coupled receptors and thioesterases. The present sequence is one  
CC such oligonucleotide. The oligonucleotides and the peptides encoded by  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate expression of the proteins listed above.  
CC Disorders that may be prevented, diagnosed and/or treated include  
CC multifactorial diseases with a genetic component, such as autoimmune  
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,  
CC leukaemia), diseases of the nervous system and an infection of pathogenic  
CC organisms  
SQ Sequence 50 BP; 7 A; 11 C; 23 G; 9 T; 0 U; 0 Other;  
Query Match 3.1%; Score 20.8; DB 4; Length 50;  
Best Local Similarity 70.0%; Pred. No. 1.7e+05;  
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 73 CTTGGGGGCCCCAGCTTCTCCAGAGGTGGCAGCAATGGCC 112  
DB 7 CGTGTGGGCTCAGTGGGTCCATAGGACAGCACTGGTC 46  
RESULT 9  
ABZ49057/c  
ID ABZ49057 standard; DNA; 40 BP.  
XX  
AC ABZ49057;  
XX  
DT 26-JUN-2003 (first entry)  
XX  
DE Human ALDH1L1/FTHPD gene polymorphic site, #5840.  
XX  
KM Human; drug metabolising enzyme; gene; drug metabolism; chromosome 3;  
KM polymorphic site; drug evaluation; drug screening; genotyping;  
KM genetic profiling; therapeutic customisation; adverse reaction;  
KM clinical trial; drug approval; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT variation replace(20..21,CCT)  
FT /\*tag= a  
XX  
PN WO200252044-A2.  
XX  
PD 04-JUL-2002.  
XX  
PF 27-DEC-2001; 2001WO-JP011592.  
XX  
PR 27-DEC-2000; 2000JP-00399443.  
XX 02-MAY-2001; 2001JP-00135256.  
XX 27-AUG-2001; 2001JP-00256862.  
XX  
PA (RIKE ) RIKEN KK.  
XX  
PI Nakamura Y, Sekine A, Iida A, Satto S;  
XX  
DR WPI; 2002-583571/62.  
XX  
PT Identifying individuals having a polymorphism, useful for determining the  
PT effectiveness or side effect of a drug or treatment protocol, comprises  
PT detecting at least one polymorphism in the drug metabolizing enzyme

PT nucleic acid.  
XX  
XX Claim 23; Page 180; 2785pp; English.  
PS  
XX  
CC Sequences ABZ43217-ABZ50887 represent polymorphic sites within genes  
CC encoding enzymes associated with drug metabolism. The invention relates  
CC to methods and compositions for identifying individuals who have at least  
CC one polymorphism in such drug metabolising enzyme-encoding genes. The  
CC polymorphisms may be identified in a nucleic acid sample using probes or  
CC primers specific for a sequence selected from ABZ43217-ABZ50887 using a  
CC variety of detection assays, including hybridisation assays, nucleic acid  
CC arrays and PCR-based methods. The invention also encompasses methods of  
CC evaluating and screening drugs using genetic polymorphism data. Genetic  
CC polymorphism data, particularly that relating to single nucleotide  
CC polymorphisms (SNPs), may be used in studying the relationship between  
CC DNA sequence variations and human diseases, conditions, and responses to  
CC drugs. SNPs are also useful as polymorphism markers for discovering genes  
CC that cause or exacerbate certain diseases. SNPs are particularly useful  
CC in the above respects as they are stable in populations, occur  
CC frequently, and have lower mutation rates than other genome variations  
CC such as repeating sequences. The detection and analysis of polymorphisms  
CC in genes encoding drug metabolising enzymes allows the customisation of  
CC drug therapies based upon the genetic profile of individual patients.  
CC This would not only take the guesswork out of selecting the drug with the  
CC greatest therapeutic effect for a particular patient, but would also  
CC reduce the likelihood of adverse reactions, thereby increasing safety.  
CC Methods of the invention are also useful in the drug discovery and  
CC approval processes. For example, individuals could be selected for  
CC clinical trials only if their genetic profiles indicate that they are  
CC capable of responding to a particular drug or drug class, and previously  
CC failed drug candidates could be revived if they were matched with more  
CC appropriate patient populations. The methods, data and compositions of  
CC the invention may therefore lead to an increase in the range of  
CC possible drug targets and decreases in the number of adverse drug  
CC reactions, failed drug trials, the time taken for a drug to be approved,  
CC the length of time patients are on medication and the number of different  
CC medications a patient needs to take before finding an effective therapy  
XX  
SQ Sequence 40 BP; 6 A; 15 C; 9 G; 10 T; 0 U; 0 Other;  
Query Match 3.1%; Score 20.6; DB 6; Length 40;  
Best Local Similarity 74.3%; Pred. No. 1.8e+05;  
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 133 CAGGAGCAGCTTCAGAAAGTTGCTGAAGTTGTGG 167  
DB 36 CAGCAGCAAGCTGGAGAGAGAGCTGCAGATTGTGG 2  
RESULT 10  
ABZ46527/c  
ID ABZ46527 standard; DNA; 40 BP.  
XX  
AC ABZ46527;  
XX  
DT 26-JUN-2003 (first entry)  
XX  
DE Human ALDH1L1/FTHPD gene polymorphic site, #3311.  
XX  
KM Human; drug metabolising enzyme; gene; drug metabolism; chromosome 3;  
KM polymorphic site; drug evaluation; drug screening; genotyping;  
KM genetic profiling; therapeutic customisation; adverse reaction;  
KM clinical trial; drug approval; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT variation replace(20..21,CCT)  
FT /\*tag= a  
XX  
PN WO200252044-A2.  
XX  
PD 04-JUL-2002.

XX 27-DEC-2001; 2001WO-JP011592.  
PF 27-DEC-2000; 2000JP-00399443.  
XX 02-MAY-2001; 2001JP-00135256.  
PR 27-AUG-2001; 2001JP-00256862.  
XX  
XX (RIKE) RIKEN KK.  
XX  
XX Nakamura Y, Sekine A, Iida A, Saito S;  
PI WPI; 2002-563571/62.  
XX  
XX Identifying individuals having a polymorphism, useful for determining the  
PT effectiveness or side effect of a drug or treatment protocol, comprises  
PT detecting at least one polymorphism in the drug metabolizing enzyme  
XX nucleic acid.  
XX  
XX Claim 23; Page 119; 2785pp; English.  
XX  
XX Sequences ABZ43217-ABZ50887 represent polymorphic sites within genes  
CC encoding enzymes associated with drug metabolism. The invention relates  
CC to methods and compositions for identifying individuals who have at least  
CC one polymorphism in such drug metabolizing enzyme-encoding genes. The  
CC polymorphisms may be identified in a nucleic acid sample using probes or  
CC primers specific for a sequence selected from ABZ43217-ABZ50887 using a  
CC variety of detection assays, including hybridisation assays, nucleic acid  
CC arrays and PCR-based methods. The invention also encompasses methods of  
CC evaluating and screening drugs using genetic polymorphism data. Genetic  
CC polymorphism data, particularly that relating to single nucleotide  
CC polymorphisms (SNPs), may be used in studying the relationship between  
CC DNA sequence variations and human diseases, conditions, and responses to  
CC drugs. SNPs are also useful as polymorphism markers for discovering genes  
CC that cause or exacerbate certain diseases. SNPs are particularly useful  
CC in the above respects as they are stable in populations, occur  
CC frequently, and have lower mutation rates than other genome variations  
CC such as repeating sequences. The detection and analysis of polymorphisms  
CC in genes encoding drug metabolizing enzymes allows the customisation of  
CC drug therapies based upon the genetic profile of individual patients.  
CC This would not only take the guesswork out of selecting the drug with the  
CC greatest therapeutic effect for a particular patient, but would also  
CC reduce the likelihood of adverse reactions, thereby increasing safety.  
CC Methods of the invention are also useful in the drug discovery and  
CC approval processes. For example, individuals could be selected for  
CC clinical trials only if their genetic profiles indicate that they are  
CC capable of responding to a particular drug or drug class, and previously  
CC failed drug candidates could be revived if they were matched with more  
CC appropriate patient populations. The methods, data and compositions of  
CC the invention may therefore lead to an increase in the range of  
CC possible drug targets and decreases in the number of adverse drug  
CC reactions, failed drug trials, the time taken for a drug to be approved,  
CC the length of time patients are on medication and the number of different  
CC medications a patient needs to take before finding an effective therapy  
XX  
XX Sequence 40 BP; 6 A; 15 C; 9 G; 10 T; 0 U; 0 Other;  
SQ  
Query Match 3.1%; Score 20.6; DB 6; Length 40;  
Best Local Similarity 74.3%; Pred. No. 1.8e+05;  
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 133 CAGGACGCGTTCAGAAAGTGTCTGAAGTTGTGG 167  
DB 36 CAGCGAGCAAGCTCGGAGAGTAGTCTGCAGATTGTGG 2  
RESULT 11  
AAI9312  
ID AAI9312 standard; DNA; 41 BP.  
XX  
XX AAI9312;  
AC  
XX 17-MAY-1999 (first entry)  
DX  
XX

DE Human granulocyte colony-stimulating factor sense primer oligo 12.  
XX  
XX Human; granulocyte colony-stimulating factor; G-CSF; hg-CSF;  
KM haematopoiesis disorder; primer; ss.  
XX  
XX Synthetic.  
OS Homo sapiens.  
XX  
XX WO9853072-A1.  
XX  
XX 26-NOV-1998.  
XX  
XX 21-MAY-1998; 98WO-KR000125.  
XX  
XX 22-MAY-1997; 97KR-00020054.  
XX  
XX (KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.  
PA (JEIL-) JEIL PHARM CO LTD.  
XX  
XX Jun H, Choi B, Lee Y, Sohn M;  
PI WPI; 1999-131691/11.  
XX  
XX New human granulocyte colony-stimulating factor (hg-CSF) containing a  
PT peptide at the N-terminus - useful in the treatment of haematopoiesis  
PT disorders.  
XX  
XX Example 1; Fig 1; 19pp; English.  
XX  
XX The present invention describes the N-terminal peptide of human  
CC granulocyte colony-stimulating factor (hg-CSF). The present invention  
CC also describes: (1) recombinant plasmid pYHM-G-CSF containing the cDNA  
CC for the peptide; and (2) E. coli BL21 (pYHM-G-CSF) (KCTC 0477BP)  
CC transformed with the plasmid. The plasmid and E. coli strain are used to  
CC produce high yields of highly purified hg-CSF, which when refolded has  
CC comparable biological activity of G-CSF. hg-CSF is used to treat  
CC haematopoietic disorders. The method of producing hg-CSF is less time-  
CC consuming and, therefore, more economical than previous methods because  
CC the protein is expressed in high yields without the need for further  
CC processing steps. AAI9290 to AAI9315 represent primer used in the  
CC synthesis of hg-CSF from an example of the present invention  
XX  
XX Sequence 41 BP; 4 A; 12 C; 13 G; 12 T; 0 U; 0 Other;  
SQ  
Query Match 3.1%; Score 20.6; DB 2; Length 41;  
Best Local Similarity 74.3%; Pred. No. 1.8e+05;  
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 498 GTGGCTGCGCTCGCTCCCTGACGCGAGCGTCTGG 532  
DB 2 GTGTCTGTGTTGCTCTCCATCTGAGAGCTTCTGG 36  
RESULT 12  
AAI96438/C  
ID AAI96438 standard; DNA; 42 BP.  
XX  
XX AAI96438;  
AC  
XX 25-JUN-2001 (first entry)  
DX  
XX Primer 172del-R.  
DE  
XX PCR primer; male sterile plant; RNase inhibitor; ss.  
KM  
XX Unidentified.  
OS  
XX WO200124616-A1.  
XX  
XX 12-APR-2001.  
XX  
XX 12-SEP-2000; 2000WO-JP006222.  
PF  
XX





04-JUL-2002.  
27-DEC-2001; 2001WO-JP011592.  
27-DEC-2000; 2000JP-00399443.  
02-MAY-2001; 2001JP-00135256.  
27-AUG-2001; 2001JP-00256862.  
(RIKE ) RIKEN KK.  
Nakamura Y, Sekine A, Iida A, Saito S;  
WPI; 2002-583571/62.  
Identifying individuals having a polymorphism, useful for determining the effectiveness or side effect of a drug or treatment protocol, comprises detecting at least one polymorphism in the drug metabolizing enzyme nucleic acid.  
Claim 23; Page 66; 2785pp; English.  
Sequences AB243217-AB250887 represent polymorphic sites within genes encoding enzymes associated with drug metabolism. The invention relates to methods and compositions for identifying individuals who have at least one polymorphism in such drug metabolizing enzyme-encoding genes. The polymorphisms may be identified in a nucleic acid sample using probes or primers specific for a sequence selected from AB243217-AB250887, using a variety of detection assays, including hybridisation assays, nucleic acid arrays and PCR-based methods. The invention also encompasses methods of evaluating and screening drugs using genetic polymorphism data. Genetic polymorphism data, particularly that relating to single nucleotide polymorphisms (SNPs), may be used in studying the relationship between DNA sequence variations and human diseases, conditions, and responses to drugs. SNPs are also useful as polymorphism markers for discovering genes that cause or exacerbate certain diseases. SNPs are particularly useful in the above respects as they are stable in populations, occur frequently, and have lower mutation rates than other genome variations such as repeating sequences. The detection and analysis of polymorphisms in genes encoding drug metabolizing enzymes allows the customisation of drug therapies based upon the genetic profile of individual patients. This would not only take the guesswork out of selecting the drug with the greatest therapeutic effect for a particular patient, but would also reduce the likelihood of adverse reactions, thereby increasing safety. Methods of the invention are also useful in the drug discovery and approval processes. For example, individuals could be selected for clinical trials only if their genetic profiles indicate that they are capable of responding to a particular drug or drug class, and previously failed drug candidates could be revived if they were matched with more appropriate patient populations. The methods, data and compositions of the invention may therefore lead to an increase in the range of possible drug targets and decreases in the number of adverse drug reactions, failed drug trials, the time taken for a drug to be approved, the length of time patients are on medication and the number of different medications a patient needs to take before finding an effective therapy

XX DE Human pigmentation trait-related PCR primer - SEQ ID No 121.  
XX KW Human; single nucleotide polymorphism; SNP; ss; melanocortin-1 receptor;  
KW genetic pigmentation trait; MC1R; agouti signaling protein; ASIP; race;  
XX hair colour; eye colour; forensic tool; PCR; primer.  
OS Homo sapiens.  
XX WO200297047-A2.  
XX PD 05-DEC-2002.  
XX PF 28-MAY-2002; 2002WO-US016789.  
XX PR 25-MAY-2001; 2001US-0293560P.  
PR 21-JUN-2001; 2001US-0300187P.  
PR 07-AUG-2001; 2001US-0310781P.  
PR 17-SEP-2001; 2001US-0323662P.  
PR 26-OCT-2001; 2001US-0344419P.  
PR 15-NOV-2001; 2001US-0334674P.  
PR 02-JAN-2002; 2002US-0346303P.  
XX (DNAP-) DNAPRINT GENOMICS INC.  
XX Prudakis T;  
XX WPI; 2003-239091/23.  
XX Inferring genetic pigmentation trait such as hair/eye color or shade from nucleic acid sample of human subject, by identifying a pigmentation-related haplotype allele of a pigmentation gene in the sample.  
XX Example 17; Page 245; 396pp; English.  
XX The invention comprises a method for inferring a genetic pigmentation trait of a human. The method involves identifying a single nucleotide polymorphism (SNP) in a pigmentation gene - where the pigmentation gene is not melanocortin-1 receptor (MC1R) and agouti signaling protein (ASIP). The method of the invention is useful for inferring a genetic pigmentation trait of a human, especially for inferring the race of a human subject. The method is useful for inferring a genetic pigmentation trait such as hair shade or colour, or eye shade or colour of a human subject. The method may be used as a forensic tool for obtaining information relating to physical characteristics of a potential crime victim or a perpetrator of a crime from a nucleic acid sample present at a crime scene. The present PCR primer is used in the exemplification of the invention  
XX SQ Sequence 45 BP; 13 A; 12 C; 15 G; 5 T; 0 U; 0 Other;  
Query Match 3.0%; Score 20; DB 7; Length 45;  
Best Local Similarity 72.2%; Pred. No. 2.5e+05;  
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 72 TCTTGGGGGCCAGCTTCTCTCCAGAGGTGGCAGCAA 107  
DB 44 TCTTGGGGCCCTGGTCCCTCCGACAGCTGCGATCGA 9  
Search completed: February 29, 2004, 12:01:11  
Job time : 308.726 secs

PD XX 04-JUL-2002.  
XX PF 27-DEC-2001; 2001WO-JP011592.  
XX KW 27-DEC-2000; 2000JP-00399443.  
PR 02-MAY-2001; 2001JP-00135256.  
PR 27-AUG-2001; 2001JP-00256862.  
XX (RIKE ) RIKEN KK.  
XX PA Nakamura Y, Sekine A, Iida A, Saito S;  
XX PI WPI; 2002-583571/62.  
XX DR Identifying individuals having a polymorphism, useful for determining the effectiveness or side effect of a drug or treatment protocol, comprises detecting at least one polymorphism in the drug metabolizing enzyme nucleic acid.  
XX PS Claim 23; Page 66; 2785pp; English.  
XX SQ Sequences AB243217-AB250887 represent polymorphic sites within genes encoding enzymes associated with drug metabolism. The invention relates to methods and compositions for identifying individuals who have at least one polymorphism in such drug metabolizing enzyme-encoding genes. The polymorphisms may be identified in a nucleic acid sample using probes or primers specific for a sequence selected from AB243217-AB250887, using a variety of detection assays, including hybridisation assays, nucleic acid arrays and PCR-based methods. The invention also encompasses methods of evaluating and screening drugs using genetic polymorphism data. Genetic polymorphism data, particularly that relating to single nucleotide polymorphisms (SNPs), may be used in studying the relationship between DNA sequence variations and human diseases, conditions, and responses to drugs. SNPs are also useful as polymorphism markers for discovering genes that cause or exacerbate certain diseases. SNPs are particularly useful in the above respects as they are stable in populations, occur frequently, and have lower mutation rates than other genome variations such as repeating sequences. The detection and analysis of polymorphisms in genes encoding drug metabolizing enzymes allows the customisation of drug therapies based upon the genetic profile of individual patients. This would not only take the guesswork out of selecting the drug with the greatest therapeutic effect for a particular patient, but would also reduce the likelihood of adverse reactions, thereby increasing safety. Methods of the invention are also useful in the drug discovery and approval processes. For example, individuals could be selected for clinical trials only if their genetic profiles indicate that they are capable of responding to a particular drug or drug class, and previously failed drug candidates could be revived if they were matched with more appropriate patient populations. The methods, data and compositions of the invention may therefore lead to an increase in the range of possible drug targets and decreases in the number of adverse drug reactions, failed drug trials, the time taken for a drug to be approved, the length of time patients are on medication and the number of different medications a patient needs to take before finding an effective therapy

QY 339 TCAGAGCTCTGATCCCGCCAGACCTGCTGGGG 374  
DB 2 TGCAGGCGCCCTGATCCCGCCAGACCAAGTTCTGGG 37  
RESULT 15  
ABT34022/c  
ID ABT34022 standard; DNA; 45 BP.  
XX AC ABT34022;  
XX XX 29-MAY-2003 (first entry)  
DT

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 29, 2004, 11:48:29 ; Search time 67.9494 Seconds  
(without alignments)  
5488.301 Million cell updates/sec

Title: US-09-904-568-1\_COPY\_132\_803  
Perfect score: 672  
Sequence: 1 agctctgtttgggggctgc.....gtgagcgcaactgcaggac 672

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 839752

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	21.2	3.2	47	4	US-09-671-317-808
C 2	20.6	3.1	47	4	US-08-671-317-639
C 3	20.2	3.0	35	4	US-09-514-247A-2
C 4	20	3.0	50	1	US-08-171-389-463
C 5	20	3.0	50	1	US-08-123-936-463
C 6	20	3.0	50	2	US-08-475-228A-463
C 7	20	3.0	50	3	US-08-482-080A-463
C 8	20	3.0	50	4	US-09-354-947-463
C 9	20	3.0	50	5	PCT-US93-12388-463
C 10	19.8	2.9	46	3	US-09-252-292-22
C 11	19.2	2.9	44	3	US-08-617-785-19
C 12	19.2	2.9	44	4	US-09-817-464-19
C 13	19.2	2.9	47	4	US-09-422-978-2676
C 14	19	2.8	27	2	US-08-859-998-80
C 15	19	2.8	27	4	US-09-325-928-80
C 16	19	2.8	27	4	US-09-325-2018-80
C 17	18.6	2.8	43	3	US-09-292-071-15
C 18	18.6	2.8	43	3	US-08-732-708C-6
C 19	18.6	2.8	43	3	US-09-292-069A-15
C 20	18.6	2.8	43	3	US-09-418-721-15
C 21	18.6	2.8	43	4	US-09-767-013-15
C 22	18.6	2.8	43	4	US-09-292-072-15
C 23	18.4	2.7	31	4	US-09-556-601-17
C 24	18.4	2.7	37	1	US-08-428-733A-7
C 25	18.4	2.7	37	1	US-08-428-733A-39
C 26	18.4	2.7	47	1	US-08-336-132-22
C 27	18.2	2.7	27	4	US-09-870-956-48

C 28	18.2	2.7	45	1	US-08-176-412-5
C 29	18.2	2.7	45	2	US-08-555-268A-5
C 30	18.2	2.7	45	2	US-08-495-695B-5
C 31	18.2	2.7	45	5	PCT-US94-14436-5
C 32	18.2	2.7	47	1	US-08-171-389-74
C 33	18.2	2.7	47	1	US-08-123-936-74
C 34	18.2	2.7	47	2	US-08-475-228A-74
C 35	18.2	2.7	47	3	US-08-482-080A-74
C 36	18.2	2.7	47	4	US-09-354-947-74
C 37	18.2	2.7	47	4	US-09-671-317-788
C 38	18.2	2.7	47	5	PCT-US93-12388-74
C 39	18	2.7	39	3	US-09-262-773-78
C 40	18	2.7	47	4	US-09-671-317-839
C 41	18	2.7	47	4	US-09-422-978-1159
C 42	18	2.7	50	2	US-08-989-394-12
C 43	18	2.7	50	3	US-09-271-365-12
C 44	18	2.7	50	4	US-09-604-013A-12
C 45	17.8	2.6	40	1	US-08-195-874-4

ALIGNMENTS

RESULT 1  
US-09-671-317-808/c

; Sequence 808, Application US/09671317  
; Patent No. 6528260

; GENERAL INFORMATION:

; APPLICANT: Blumenfeld, Marta

; APPLICANT: Chumakov, Ilya

; APPLICANT: Bougueleret, Lydie

; APPLICANT: Cohen, Annick

; TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM

; FILE REFERENCE: 62.US3.CIP

; CURRENT APPLICATION NUMBER: US/09/671,317

; CURRENT FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US 09/536,178

; PRIOR FILING DATE: 2000-03-23

; PRIOR APPLICATION NUMBER: PCT/IB00/00403

; PRIOR FILING DATE: 2000-03-24

; PRIOR APPLICATION NUMBER: US 60/126,269

; PRIOR FILING DATE: 1999-03-25

; PRIOR APPLICATION NUMBER: US 60/131,961

; PRIOR FILING DATE: 1999-04-30

; NUMBER OF SEQ ID NOS: 977

; SOFTWARE: Patent.pm

; SEQ ID NO 808

; LENGTH: 47

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE:

; NAME/KEY: allele

; LOCATION: 24

; OTHER INFORMATION: 12-132-437 : polymorphic base A or C

US-09-671-317-808

Query Match 3.2%; Score 21.2; DB 4; Length 47;

Best Local Similarity 65.9%; Pred. No. 2.3e+03;

Matches 29; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

QY 96 AGTGACGCCCTGGCTCAGCTTGAAGGACTTAGACAGGAGCA 139

DB 44 AGCGATGCCCTTGGGCGAGANTTGGAGGGGACAGGAGCA 1

RESULT 2

US-09-671-317-639

; Sequence 639, Application US/09671317

; Patent No. 6528260

; GENERAL INFORMATION:

; APPLICANT: Blumenfeld, Marta

; APPLICANT: Chumakov, Ilya

; APPLICANT: Bougueleret, Lydie

RESULT 4  
US-08-171-389-463/c  
; Sequence 463, Application US/08171389  
; Patent No. 5578444

RESULT 5  
US-08-123-936-463/c  
; Sequence 463, Application US/08123936  
; Patent No. 5726014  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Cynthia A.  
; APPLICANT: Cantor, Charles R.

APPLICANT: Andrews, Beth M.  
APPLICANT: Turin, Lisa M.  
TITLE OF INVENTION: Screening Assay for the Detection of  
DNA-Binding Molecules  
NUMBER OF SEQUENCES: 640  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genelabs Technologies, Inc.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: USA  
ZIP: 94063  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/123,936  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/996,783  
FILING DATE: 23-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/723,618  
FILING DATE: 27-JUN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 4600-0075.32/G19P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 463:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Human insulin-like growth factor II  
INDIVIDUAL ISOLATE: gene  
US-08-123-936-463

Query Match 3.0%; Score 20; DB 1; Length 50;  
Best Local Similarity 65.9%; Pred. No. 5.5e-03;  
Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
QY 259 GCGGGCACACACCTGCTCCAGCAGGCTCTCCGGCTGCCCC 302  
DB 49 GCGGGCGCCACGCGCGCTCTTATAGTCGGCGCAGCGCGGCC 6

RESULT 6  
US-08-475-228A-463/c  
Sequence 463, Application US/08475228A  
Patent No. 5869241  
GENERAL INFORMATION:  
APPLICANT: Edwards, Cynthia A.  
APPLICANT: Cantor, Charles R.  
APPLICANT: Andrews, Beth M.  
APPLICANT: Turin, Lisa M.  
APPLICANT: Fry, Kirk E.  
TITLE OF INVENTION: Sequence-Directed DNA Binding  
NUMBER OF SEQUENCES: 664  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genelabs Technologies, Inc.  
STREET: 505 Penobscot Drive  
CITY: Redwood City

STATE: CA  
COUNTRY: USA  
ZIP: 94063  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,228A  
FILING DATE: 06-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/123,936  
FILING DATE: 17-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/996,783  
FILING DATE: 23-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/723,618  
FILING DATE: 27-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/081,070  
FILING DATE: 22-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Stratford, Carol A.  
REGISTRATION NUMBER: 34,444  
REFERENCE/DOCKET NUMBER: 4600-0175.21/G19P3D2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 463:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Human insulin-like growth factor II  
INDIVIDUAL ISOLATE: gene  
US-08-475-228A-463

Query Match 3.0%; Score 20; DB 2; Length 50;  
Best Local Similarity 65.9%; Pred. No. 5.5e-03;  
Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
QY 259 GCGGGCACACACCTGCTCCAGCAGGCTCTCCGGCTGCCCC 302  
DB 49 GCGGGCGCCACGCGCGCTCTTATAGTCGGCGCAGCGCGGCC 6

RESULT 7  
US-08-482-080A-463/c  
Sequence 463, Application US/08482080A  
Patent No. 6010849  
GENERAL INFORMATION:  
APPLICANT: Edwards, Cynthia A.  
APPLICANT: Cantor, Charles R.  
APPLICANT: Andrews, Beth M.  
APPLICANT: Turin, Lisa M.  
APPLICANT: Fry, Kirk E.  
TITLE OF INVENTION: Sequence-Directed DNA Binding  
NUMBER OF SEQUENCES: 664  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genelabs Technologies, Inc.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: USA  
ZIP: 94063  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,080A  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/482,080  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/171,389  
FILING DATE: 20-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/123,936  
FILING DATE: 17-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/996,783  
FILING DATE: 23-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/723,618  
FILING DATE: 27-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/081,070  
FILING DATE: 22-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Brady, John F.  
REGISTRATION NUMBER: 39,118  
REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 324-0880  
TELEFAX: (650) 324-0960  
INFORMATION FOR SEQ ID NO: 463:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Human insulin-like growth factor II  
INDIVIDUAL ISOLATE: gene  
US-08-482-080A-463

Query Match 3.0%; Score 20; DB 3; Length 50;  
Best Local Similarity 65.9%; Pred. No. 5.5e+03;  
Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 259 GGGGCGACACACACCTGCTCCAGCAGGCTCTCCGGGTGCCCC 302  
Db 49 GGGGCGCCACCGCCCGGCTCTTATAGTCGCGCCAGCGCGGCC 6

RESULT 8  
US-09-354-947-463/c  
Sequence 463, Application US/09354947  
Patent No. 6384208  
GENERAL INFORMATION:  
APPLICANT: Edwards, Cynthia A.  
APPLICANT: Cantor, Charles R.  
APPLICANT: Andrews, Beth M.  
APPLICANT: Turin, Lisa M.  
APPLICANT: Fy, Kirk E.  
TITLE OF INVENTION: Sequence-Directed DNA Binding  
TITLE OF INVENTION: Molecules, Compositions and Methods  
NUMBER OF SEQUENCES: 664  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genelabs Technologies, Inc.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: USA  
ZIP: 94063  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/354,947  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/482,080  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/171,389  
FILING DATE: 20-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/123,936  
FILING DATE: 17-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/996,783  
FILING DATE: 23-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/723,618  
FILING DATE: 27-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/081,070  
FILING DATE: 22-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Brady, John F.  
REGISTRATION NUMBER: 39,118  
REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 324-0880  
TELEFAX: (650) 324-0960  
INFORMATION FOR SEQ ID NO: 463:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Human insulin-like growth factor II  
INDIVIDUAL ISOLATE: gene  
US-09-354-947-463

Query Match 3.0%; Score 20; DB 4; Length 50;  
Best Local Similarity 65.9%; Pred. No. 5.5e+03;  
Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 259 GGGGCGACACACACCTGCTCCAGCAGGCTCTCCGGGTGCCCC 302  
Db 49 GGGGCGCCACCGCCCGGCTCTTATAGTCGCGCCAGCGCGGCC 6

RESULT 9  
PCT-US93-12388-463/c  
Sequence 463, Application PC/TUS9312388  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Sequence-Directed DNA Binding  
TITLE OF INVENTION: Molecules, Compositions and Methods  
NUMBER OF SEQUENCES: 641  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genelabs Technologies, Inc.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: USA  
ZIP: 94063  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/12388  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/123,936  
FILING DATE: 17-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/996,783  
FILING DATE: 23-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 4600-0175.41/G19PCT2  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 463:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Human insulin-like growth factor II  
INDIVIDUAL ISOLATE: gene  
PCT-US93-12388-463

Query Match 3.0%; Score 20; DB 5; Length 50;  
Best Local Similarity 65.9%; Pred. No. 5.5e+03;  
Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
QY 259 GCGGCGCCACACACCCCTGCTCCAGCAGGCTCTCCGGCTGCCCC 302  
|||||  
DB 49 GCGGCGCCACACACCCCTGCTCCAGCAGGCTCTCCGGCTGCCCC 6

RESULT 10  
US-09-252-292-22  
Sequence 22, Application US/09252292B  
Patent No. 6245968  
GENERAL INFORMATION:  
APPLICANT: Boudec, Phillipe  
APPLICANT: Rodgers, Matthew  
APPLICANT: Dumas, Florence  
TITLE OF INVENTION: Mutated hydroxyphenylpyruvate dioxygenase, DNA  
TITLE OF INVENTION: sequence and isolation of plants which contain such a  
TITLE OF INVENTION: gene and which are tolerant to herbicides  
FILE REFERENCE: 5500731  
CURRENT APPLICATION NUMBER: US/09/252,292B  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 08/982,772  
PRIOR FILING DATE: 1997-12-02  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 22  
LENGTH: 46  
TYPE: DNA  
ORGANISM: Plant  
US-09-252-292-22

Query Match 2.9%; Score 19.8; DB 3; Length 46;  
Best Local Similarity 69.2%; Pred. No. 6.1e+03;  
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 550 GATGACTGACACACCGCTTCGAGTGCCCATAGCCAAA 588  
|||||  
DB 4 GATGACTGACACACCGCTTCGAGTGCCCATAGCCAAA 42

RESULT 11  
US-08-617-785-19/c

Sequence 19, Application US/08617785E  
Patent No. 6228610  
GENERAL INFORMATION:  
APPLICANT: Flor, Peter J.  
APPLICANT: Kuhn, Ranier  
APPLICANT: Lindaur, Kristen  
APPLICANT: Puttner, Irene  
APPLICANT: Knopfel, Thomas  
TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,  
FILE REFERENCE: 4-19679/A/PCT  
CURRENT APPLICATION NUMBER: US/08/617,785E  
CURRENT FILING DATE: 1996-03-19  
EARLIER APPLICATION NUMBER: PCT/EP94/02991  
EARLIER FILING DATE: 1994-09-07  
EARLIER APPLICATION NUMBER: EPO 9416553.7  
EARLIER FILING DATE: 1994-08-19  
EARLIER APPLICATION NUMBER: EPO 93810663.0  
EARLIER FILING DATE: 1993-09-20  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 19  
LENGTH: 44  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:  
OTHER INFORMATION: Oligonucleotide primer  
US-08-617-785-19

Query Match 2.9%; Score 19.2; DB 3; Length 44;  
Best Local Similarity 67.5%; Pred. No. 9.2e+03;  
Matches 27; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
QY 266 CACACACCTGCTCCAGCAGGCTCTCCGCTGCCCTCC 305  
|||||  
DB 42 CCGCTCTAGCCCTAGCAGGCTCCGCTGCCCTCCGCTCC 3

RESULT 12  
US-09-817-464-19/c  
Sequence 19, Application US/09817464  
Patent No. 6515107  
GENERAL INFORMATION:  
APPLICANT: Flor, Peter J.  
APPLICANT: Kuhn, Ranier  
APPLICANT: Lindaur, Kristen  
APPLICANT: Puttner, Irene  
APPLICANT: Knopfel, Thomas  
TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,  
FILE REFERENCE: 4-19679/A/PCT  
CURRENT APPLICATION NUMBER: US/09/817,464  
CURRENT FILING DATE: 2001-03-26  
EARLIER APPLICATION NUMBER: US/08/617,785  
EARLIER FILING DATE: 1996-03-19  
EARLIER APPLICATION NUMBER: EPO 9416553.7  
EARLIER FILING DATE: 1994-08-19  
EARLIER APPLICATION NUMBER: EPO 93810663.0  
EARLIER FILING DATE: 1993-09-20  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 19  
LENGTH: 44  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION:  
OTHER INFORMATION: Oligonucleotide primer  
US-09-817-464-19

Query Match 2.9%; Score 19.2; DB 4; Length 44;  
Best Local Similarity 67.5%; Pred. No. 9.2e+03;

```
Matches 27; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 266 CACACACCTGTCAGAGGCTCCGGCTGCCCGCTGCCCTGC 305
Db 42 CCGGTCCTAGCCTAGCAGGCGCTCGGCGCCCTGCAGC 3

RESULT 13
US-09-422-978-2676
; Sequence 2676, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSET.020CP1
; CURRENT APPLICATION NUMBER: US/09/422,978
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/298,850
; EARLIER FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,732
; EARLIER FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: US 60/082,614
; EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 2676
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 99-13864-64 : polymorphic base G or T
US-09-422-978-2676

Query Match 2.9%; Score 19.2; DB 4; Length 47;
Best Local Similarity 70.6%; Pred. No. 9.5e+03;
Matches 24; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 272 CCTGCTCAGCAGGCTCTCCGGCTGCCCGCTGCCCTGC 305
Db 10 CCCTCCTCCACAGCTGCCCTCCTCCCGCTGC 43

RESULT 14
US-08-859-998-80/c
; Sequence 80, Application US/08859998
; Patent No. 5994076
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Jekhade, George
; APPLICANT: Bibilashvili, Robert
; TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1375
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: CA
; COUNTRY: US
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/859,998
; FILING DATE: 21-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; REGISTRATION NUMBER: 08/859,998
; REFERENCE/DOCKET NUMBER: 08/859,998
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
```

```
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Field, Bret E.
REGISTRATION NUMBER: 37,620
REFERENCE/DOCKET NUMBER: 09096/002001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
OTHER INFORMATION: oligonucleotide primer
US-08-859-998-80

Query Match 2.8%; Score 19; DB 2; Length 27;
Best Local Similarity 81.5%; Pred. No. 8.3e+03;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 43 GCTGACAGTCACAGTGCGCGGTCAGT 69
Db 27 GCAGACAGTCACACTGGTTGGTCAGT 1

RESULT 15
US-09-225-328-80/c
; Sequence 80, Application US/09225928
; Patent No. 6352829
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Jekhade, George
; APPLICANT: Bibilashvili, Robert
; TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1375
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: CA
; COUNTRY: US
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/225,928
; FILING DATE: 05-Jan-1999
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/859,998
; FILING DATE: 21-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Field, Bret E.
; REGISTRATION NUMBER: 37,620
; REFERENCE/DOCKET NUMBER: 09096/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
```



```

;
; FEATURE:
;
; OTHER INFORMATION: oligonucleotide primer
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 80:
US-09-225-928-80

```

Query Match 2.8%; Score 19; DB 4; Length 27;  
Best Local Similarity 81.5%; Pred. No. 8.3e+03;  
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 43 GGTGACAGTACACAGTGGCCGGTCAGT 69  
Db 27 GCAGACAGTACACACTGGTTTGGTCAGT 1

Search completed: February 29, 2004, 14:42:39  
Job time : 70.9494 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: February 29, 2004, 09:53:57 ; Search time 303.274 Seconds  
(without alignments)  
9413.223 Million cell updates/sec

Title: US-09-904-568-1\_COPY\_132\_803  
Perfect score: 672  
Sequence: 1 agtgtctttgggggtgc.....gtgagcgcaactgcagcagc 672

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 3185356

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	21.8	3.2	49	6	ABA99919 Human TNF
C 2	21.4	3.2	50	7	ABZ24835 Herpes S1
C 3	21	3.1	30	3	AAA71444 Human meg
C 4	21	3.1	33	6	ABL40388 Primer 3
C 5	21	3.1	40	2	AAT89030 Solid pha
C 6	20.8	3.1	42	2	AAQ52858 Cytomegal
C 7	20.8	3.1	49	2	AAT80503 Hepatoma
C 8	20.2	3.0	35	2	AAI66427 Kringl p
C 9	20.2	3.0	41	4	AAI66427 Kringl p
C 10	20.2	3.0	44	2	AAQ22113 S3 chim
C 11	20.2	3.0	50	4	AAI34545 Human SNP
C 12	20	3.0	50	2	AAQ69713 Human ins
C 13	20	3.0	50	2	AAT64175 Human ins
C 14	20	3.0	50	2	AAI17463 Test sequ
C 15	20	3.0	50	6	ABK52954 DNA bindi
C 16	20	3.0	50	6	ABZ00686 Human leu
C 17	20	3.0	50	10	AD80493 Duplex ol
C 18	19.6	2.9	40	7	ACA55333 Human mod
C 19	19.6	2.9	47	3	AAZ68329 Human map
C 20	19.6	2.9	50	8	ACD17187 Mouse ant
C 21	19.4	2.9	41	6	ABZ43700 Human HNK
C 22	19.4	2.9	41	6	ABZ49884 Human HNK
C 23	19.4	2.9	41	6	ABV75966 Human ATP

24	19.4	2.9	50	4	AAH89679 Human rib
C 25	19.4	2.9	50	6	ABZ00318 Human leu
C 26	19.2	2.9	32	2	AAV90578 Reverse P
C 27	19.2	2.9	33	6	ABK96448 PCR prime
C 28	19.2	2.9	50	4	AAI29231 Human SNP
C 29	19.2	2.9	50	6	ABZ01371 Human leu
C 30	19.2	2.9	50	6	ABZ01296 Human leu
C 31	19	2.8	27	6	ABK65992 Human gen
C 32	19	2.8	33	7	ABQ82800 Human Akt
C 33	19	2.8	33	7	ABV76886 3' PCR pr
C 34	19	2.8	33	7	ABQ82813 Human Akt
C 35	19	2.8	33	7	ABA00519 Human Akt
C 36	19	2.8	33	7	AAI47158 Human Akt
C 37	19	2.8	33	9	ADD96154 Human Akt
C 38	19	2.8	33	9	ADE06642 Primer #4
C 39	19	2.8	33	9	ADE15856 Primer #4
C 40	19	2.8	39	6	ABA98210 Oligonuc
C 41	19	2.8	46	6	ABK52805 Human pro
C 42	19	2.8	50	4	AAI30792 Human SNP
C 43	18.8	2.8	39	2	AAI61009 L-galacto
C 44	18.8	2.8	40	2	AAV51259 Maize pol
C 45	18.8	2.8	42	4	AAI02645 Human CGO

ALIGNMENTS

RESULT 1  
ID ABA99919/c  
XX ABA99919 standard; DNA; 49 BP.  
AC ABA99919;  
XX  
DT 05-JUL-2002 (first entry)  
XX Human TNF PCR primer SEQ ID 13.  
DE  
XX Prodrug; TNF; tumour necrosis factor; selectokine; chimeric; W24; W33;  
KW cytosolic; immunomodulatory; antiangiogenic; apoptosis inducer; tenascin;  
KW gene therapy; scfv antibody O84; fibroblast activation protein; tenascin;  
KW solid tumour; angiogenesis; treatment; infection; metabolic disease; PCR;  
KW primer; ss.  
XX Homo sapiens.  
OS  
XX WO200222833-Al.  
PN  
XX 21-MAR-2002.  
PD  
XX 17-SEP-2001; 2001WO-EP010730.  
PF  
XX 15-SEP-2000; 2000DE-01045592.  
PR  
XX (UYST-) UNIV STUTTGART.  
PA (PFIZ/) PFIZENMAIER K.  
XX Pfizermaier K, Wuest T, Moosmayer D, Grell M, Scheurich P;  
PI WPI; 2002-362351/39.  
XX New polypeptide prodrug, useful e.g. for treating tumors, contains  
DR targeting region, active agent and attached inhibitor that is  
XX proteolytically cleaved in target cells.  
PT Example 6; Page 47; 52pp; German.  
PS This invention describes a novel polypeptide (I) comprising, in the N to  
XX C direction, a region (R1) that recognises selectively a specific  
CC macromolecule on a cell surface and/or a component of the extracellular  
CC matrix, peptide linker, a region (R2) with biological activity for a  
CC specific target molecule, a region (R3) that has a processing site and a  
CC region (R4) that inhibits the activity of R2, by intramolecular bonding  
CC and/or interaction. The products of the invention have cytostatic,

CC immunomodulatory and antiangiogenic activity, induce apoptosis and can be  
CC used for gene therapy. Kym-1 cells (20000) were incubated with the  
CC prodrug W24, containing, essentially, the single-chain Fv antibody O54,  
CC specific for human fibroblast activation protein, trimerization linker,  
CC mutant form of the tumour necrosis factor (TNF) precursor protein, a  
CC region with a proteolytic cleavage site, and human TNF receptor-1  
CC fragment, and with trypsin (activator) for 5 minutes. After 16 hours,  
CC cell viability was determined by MTT staining. Activated W24 had LD50  
CC about 0.5 ng/ml, comparable with that for wild-type TNF and 4000 times  
CC higher than for uncleaved W24. (1), also nucleic acids encoding them and  
CC related vectors, are useful particularly for treating solid tumours  
CC and/or pathological angiogenesis, also generally for treating infections  
CC and metabolic diseases. (1) are prodrug forms of R2 that have  
CC unacceptable toxicity when administered systemically (specifically tumour  
CC necrosis factor) and allow these compounds to be administered safely with  
CC retention of, or even increase in, therapeutic activity. R2 is released  
CC only in target tissue, resulting in a high local concentration, and  
CC activity is potentiated by co-activation of receptors. This sequence  
CC represents a PCR primer for the amplification of the human TNF fragment  
CC used in the construction of the TNF-selectokine W24 and W33 prodrugs  
CC described in the disclosure of the invention  
XX  
SQ Sequence 49 BP; 12 A; 15 C; 12 G; 10 T; 0 U; 0 Other;

Query Match 3.2%; Score 21.8; DB 6; Length 49;  
Best Local Similarity 70.7%; Pred. No. 2.6e+04;  
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 74 TGGGTCCAGCCTCTCCAGAGTCAGCGCGCTGCTCAG 114  
|||||  
Db 46 TGGGATCATTGCTCTCCGAATCGAGGTCGTGGATCCG 6

## RESULT 2

ABZ24835  
ID ABZ24835 standard; DNA; 50 BP.

XX  
AC ABZ24835;

XX  
DT 03-MAR-2003 (first entry)

XX Herpes Simplex Virus thymidine kinase PCR primer TK3'-1.

XX PCR; primer; intranuclear receptor; sex hormone receptor;  
KW androgen receptor; AR; oestrogen receptor; ER; ss.

XX Herpes Simplex Virus.

XX JP2002247986-A.

XX 03-SEP-2002.

XX 23-FEB-2001; 2001JP-00048876.

XX 23-FEB-2001; 2001JP-00048876.

XX (MITA) MITSUI CHEM INC.

XX WPI; 2003-096457/09.

XX Cells for evaluation of agonistic and antagonistic activities to a  
PT nuclear receptor of a substance, comprise a stable expression cell line  
PT of an intranuclear receptor gene and a reporter gene.

XX Example 4; Page 11; 33pp; Japanese.

XX The present invention relates to cells obtained by introduction of a  
CC reporter expression nucleic acid, which comprises a response element and  
CC a promoter downstream of the response element. The promoter controls  
CC transcription in animal host cells to significantly stimulate the  
CC expression of the reporter gene in the presence of a ligand bound to an  
CC intranuclear receptor. The cells are used for evaluation of a sex hormone  
CC receptor e.g. androgen receptor (AR) and oestrogen receptor (ER) and

CC allow sensitive and correct determination of sex hormonal activity. The  
CC present sequence is a PCR primer, which was used in an example from the  
CC invention

XX Sequence 50 BP; 7 A; 12 C; 18 G; 13 T; 0 U; 0 Other;

Query Match 3.2%; Score 21.4; DB 7; Length 50;  
Best Local Similarity 66.0%; Pred. No. 3.4e-04;

XX Matches 31; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 136 AGCACCTTCAGAAAGTTGTGAACTTGTAGTCGGGCGCTGCATGG 182

Db 4 AGCTTCTGGGCACGCTGTTGACGCTGTAGCGGGTCTGCTGCAGGG 50

## RESULT 3

AAA71444/c  
ID AAA71444 standard; DNA; 30 BP.

XX AAA71444;

XX 01-DEC-2000 (first entry)

XX Human megasin promoter PCR primer SEQ ID NO: 11.

XX Promoter; megasin; human; protein isolation; screening. PCR primer; ss.

XX Homo sapiens.

XX WO2000043528-A1.

XX 27-JUL-2000.

XX 25-JAN-2000; 2000WO-JP000350.

XX 25-JAN-1999; 99JP-00015667.

XX (KURO/) KUROKAWA K.

XX (MIYA/) MIYATA T.

XX Miyata T;

XX WPI; 2000-543257/49.

XX DNA for promoter region of megasin useful for screening proteins.

XX Example 5; Page 38; 45pp; Japanese.

XX This invention describes a novel DNA sequence (1) representing a promoter  
CC region having part or all of a specific base sequence. The invention also  
CC describes (1) a vector containing (1); (2) a cell transformed by the  
CC above vector; and (3) protein produced using (1). (1) is useful for  
CC screening and isolating proteins (especially transcription factors).  
CC AAA71434-A71469 represent PCR primers used in the method described in the  
CC invention

XX Sequence 30 BP; 6 A; 10 C; 5 G; 9 T; 0 U; 0 Other;

Query Match 3.1%; Score 21; DB 3; Length 30;  
Best Local Similarity 82.8%; Pred. No. 3.7e+04;

XX Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 135 GAGCACCTTCAGAAAGTTGTGAACTTG 163

Db 30 GAGCACCTTCAGATAGAGCTGAACCTTG 2

## RESULT 4

ABL40388/c  
ID ABL40388 standard; cDNA; 33 BP.

XX ABL40388;

XX

DT 01-JUL-2002 (first entry)  
DE Primer 3 relative to HOMO phosphoryl transport protein 12.  
XX HOMO phosphoryl transport protein 12; cancer; haemopathy; phlogosis;  
XX immune disease; HIV; human immunodeficiency virus; cytostatic;  
XX haemostatic; virucide; immunomodulatory; antiinflammatory; inflammation;  
XX gene therapy; malignant tumour; PCR; primer; ss.  
XX Unidentified.  
XX W0200200698-A1.  
XX 03-JAN-2002.  
XX 14-MAY-2001; 2001WO-CN000792.  
XX 16-MAY-2000; 2000CN-00115689.  
XX (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.  
XX Mao Y, Xie Y;  
XX WPI; 2002-075663/10.  
XX Polypeptide-HOMO phosphoryl transport protein 12 and encoding  
XX polynucleotide, used in diagnosis and treatment of malignant tumors,  
XX haemopathy, human immunodeficiency virus infection, immunological diseases  
XX and inflammation.  
XX Example 4; Page 18; 37pp; Chinese.  
XX The invention relates to an isolated polypeptide of polypeptide-HOMO  
XX phosphoryl transport protein 12. The activity of the polypeptide of the  
XX invention may be described as, cytostatic, haemostatic, virucide,  
XX immunomodulatory and antiinflammatory. Polypeptides and polynucleotides  
XX of the invention are used in diagnosis and treatment of malignant tumour,  
XX haemopathy, human immunodeficiency virus (HIV) infection, immunological  
XX diseases various inflammations, and phlogosis. They may also be used in  
XX gene therapy. The current sequence represents a primer relative to the  
XX HOMO phosphoryl transport protein 12 of the invention  
XX Sequence 33 BP; 5 A; 9 C; 13 G; 6 T; 0 U; 0 Other;  
Query Match 3.1%; Score 21; DB 6; Length 33;  
Best Local Similarity 82.8%; Pred. No. 3.8e+04;  
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 280 CACGAGCTCTCCGGCTGCGCCCTGCTAG 308  
DB 33 CACGAGCTCTCAGGGCGCCCATGCTAG 5  
RESULT 5  
AAT89030  
ID AAT89030 standard; DNA; 40 BP.  
XX AAT89030;  
XX 04-FEB-1998 (first entry)  
XX Solid phase gene HPA-2 DNA probe.  
XX Solid phase gene; synthetic gene; gene synthesis; trityl group;  
XX hydrophobic carrier; PCR primer; detection probe; HPA; ss.  
XX Synthetic.  
XX JP09257798-A.  
XX 03-OCT-1997.  
XX 19-MAR-1996; 96JP-00062885.  
XX

XX 19-MAR-1996; 96JP-00062885.  
XX (SUMQ) SUMITOMO METAL IND LTD.  
XX WPI; 1997-540228/50.  
XX Solid phase gene - useful for detection of genes in a sample.  
XX Example 1; Page 9; 12pp; Japanese.  
XX The present sequence represents a DNA probe which is used to demonstrate  
XX a new method for the detection of genes, using a synthetic gene in the  
XX solid phase. The gene has a (substituted) trityl group which binds to a  
XX hydrophobic carrier through an end protective group. The method is simple  
XX and the solid phased gene is stable. Gene amplification was carried out  
XX using a primer set to contain a point mutation determining each antigen  
XX type in the amplification region for HPA-1 to HPA-6. A synthetic DNA was  
XX prepared and a probe was immobilised on the plate. The present probe  
XX binds to HPA-2  
XX Sequence 40 BP; 7 A; 9 C; 11 G; 13 T; 0 U; 0 Other;  
Query Match 3.1%; Score 21; DB 2; Length 40;  
Best Local Similarity 73.0%; Pred. No. 4.1e+04;  
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 308 GTCTAAAGCCAGATGCTTCCAGGAGAGAGCTCCAGGA 344  
DB 4 GTCTCAGCTCATTCCTTCAGGTAGAGCTCTTGGGA 40  
RESULT 6  
AAQ52858  
ID AAQ52858 standard; RNA; 42 BP.  
XX AAQ52858;  
XX 25-MAR-2003 (revised)  
DT 26-MAY-1994 (first entry)  
XX Cytomegalovirus target sequence 35.  
XX RNA; enzyme; enzymatic RNA molecule; ERM; cleave; RNA; mRNA; HbRNA;  
XX picornavirus; HIV; immunodeficiency virus; hepatitis B virus; HBV;  
XX papilloma virus; HPV; Epstein-Barr virus; EBV; TGLV;  
XX T-cell leukaemia virus; hepatitis C virus; HCV; cytomegalovirus;  
XX influenza virus; HSV; herpes simplex virus; vector; immune response;  
XX antibody; ribozyme; viral RNA; treatment; ss.  
XX Synthetic.  
XX WO9323569-A1.  
XX 25-NOV-1993.  
XX 29-APR-1993; 93WO-US004020.  
XX 11-MAY-1992; 92US-00882699.  
XX 14-MAY-1992; 92US-00882712.  
XX 14-MAY-1992; 92US-00882713.  
XX 14-MAY-1992; 92US-00882714.  
XX 14-MAY-1992; 92US-00882823.  
XX 14-MAY-1992; 92US-00882824.  
XX 14-MAY-1992; 92US-00882886.  
XX 14-MAY-1992; 92US-00882888.  
XX 14-MAY-1992; 92US-00882889.  
XX 14-MAY-1992; 92US-00882921.  
XX 14-MAY-1992; 92US-00882922.  
XX 14-MAY-1992; 92US-00883823.  
XX 14-MAY-1992; 92US-00883849.  
XX 14-MAY-1992; 92US-00884073.  
XX 14-MAY-1992; 92US-00884074.

```
PR 14-MAY-1992; 92US-00884333.
PR 14-MAY-1992; 92US-00884422.
PR 14-MAY-1992; 92US-00884431.
PR 14-MAY-1992; 92US-00884436.
PR 14-MAY-1992; 92US-00884521.
PR 31-JUL-1992; 92US-00923738.
PR 26-AUG-1992; 92US-00935854.
PR 26-AUG-1992; 92US-00936086.
PR 18-SEP-1992; 92US-00948359.
PR 15-OCT-1992; 92US-00963322.
PR 07-DEC-1992; 92US-00987129.
PR 07-DEC-1992; 92US-00987130.
PR 07-DEC-1992; 92US-00987133.
XX
XX (RIBO-) RIBOZYME PHARM INC.
XX
XX Draper KG, Dudycz LW, Mcswiggen JA, Macejak DG, Holecek JJ;
XX Mamone JA;
XX
XX WPI; 1993-386599/48.
XX
XX Enzymatic RNA molecules - used to inhibit viral replication, infection
XX and gene expression.
XX
XX Claim 5; Fig 13; 287pp; English.
XX
XX The sequences (AAQ52824-Q52890) are pref. Cytomegalovirus target
XX sequences for enzymatic RNA molecules. The RNA molecules are
XX complementary to a substrate binding region in the specified gene target.
XX They also have enzymatic activity, in that they specifically cleave RNA
XX in the target. The ERMs interfere with viral replication and therefore
XX have anti-viral properties. They can be used to attenuate viruses to be
XX used in vaccines. (Updated on 25-MAR-2003 to correct PN field.) (Updated
XX on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct
XX PI field.)
XX
XX SQ Sequence 42 BP; 9 A; 9 C; 17 G; 0 T; 7 U; 0 Other;
Query Match 3.1%; Score 20.8; DB 2; Length 42;
Best Local Similarity 60.0%; Pred. No. 4.8e+04;
Matches 24; Conservative 4; Mismatches 12; Indels 0; Gaps 0;
QY 201 TTGTGGAGCAACTGTGTGCAGCGCCACCTGTCTCAGAGAG 240
:|||||:|:|||||:|||||:|||||:|||||:
DB 2 UGUGGAUGACCUACGCGGCCAUCGCGGAGGAGUCAGAUGAG 41
|||||:|||||:|||||:|||||:|||||:
RESULT 7
AAT80503
AC AAT80503 standard; DNA; 49 BP.
XX
XX AAT80503;
XX
XX 31-OCT-1997 (first entry)
XX
XX Hepatoma AS-30D Type II hexokinase promoter fragment from -118.
XX
XX Response element; Z-DNA; neoplasia; hexokinase II; glycolysis; cancer;
XX gene therapy; diabetes; tumour; rat; ss.
XX
XX Rattus rattus.
XX
XX WO9704104-A2.
XX
XX 06-FEB-1997.
XX
XX 12-JUL-1996; 96WO-US011673.
XX
XX 14-JUL-1995; 95US-0001199P.
XX
XX (UYJO ) UNIV. JOHNS HOPKINS.
XX
XX Pedersen PL, Mathupala SP, Rempel A;
XX
XX The invention relates to methods for identifying potential agonists or
```

```
XX WPI; 1997-132643/12.
XX
XX New transcription regulating fragments of hexokinase II DNA contg.
XX response element - and methods for diagnosis or treatment of neoplasias
XX that over-express hexokinase II and for regulating glycolysis.
XX
XX Claim 1; Fig 11; 104pp; English.
XX
XX The present sequence represents a segment of the hepatoma AS-30D Type II
XX hexokinase promoter region. Response elements (transcription factor
XX binding site) in this fragment may consist of all or part of the present
XX sequence. AS-30D is a new isolated hexokinase II. The present DNA
XX fragment is capable of regulating transcription of a downstream open
XX reading frame and contains at least one response element. The present DNA
XX fragment may be coupled to a reporter gene and used to screen for
XX potential drugs that affect regulated transcription of tumour hexokinase
XX II. Alternatively it may be coupled to a toxic gene and used to treat
XX cells that over-express hexokinase II, such as those present in patients
XX with cancer. It may also be used in gene therapy to treat diabetes. The
XX DNA fragment increases glycolysis in cells and express homologous or
XX heterologous protein. Probes of the DNA fragment are used in the method
XX for diagnosing a neoplasia that over-expresses hexokinase. The new
XX response elements are active only in tumours, not in normal cells
XX
XX SQ Sequence 49 BP; 11 A; 21 C; 12 G; 5 T; 0 U; 0 Other;
Query Match 3.1%; Score 20.8; DB 2; Length 49;
Best Local Similarity 64.8%; Pred. No. 5e+04;
Matches 31; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 255 GCTGGGGGGCACACACACCTGTCTCCAGCAGGCTCTCCGGTGTCCCCC 302
|||||:|||||:|||||:|||||:|||||:
DB 2 GCCGGAGAGGGTACACACCTCTTCCGCGACCAATGAGCGGCCAC 49
|||||:|||||:|||||:|||||:|||||:
RESULT 8
AAX27583
ID AAX27583 standard; DNA; 35 BP.
XX
XX AAX27583;
XX
XX 27-MAY-1999 (first entry)
XX
XX PPAR cDNA amplifying primer 2.
XX
XX Peroxisome proliferator-activated receptor; PPAR; reporter gene; CAT;
XX transcriptional coupling factor; screening; CBP; SRC-1; lacZ;
XX antidiabetic; antiobesity; PCR primer; ss.
XX
XX Synthetic.
XX
XX WO9910532-A1.
XX
XX 04-MAR-1999.
XX
XX 24-AUG-1998; 98WO-JP003734.
XX
XX 27-AUG-1997; 97JP-00231084.
XX
XX (TANA ) TANABE SEIYAKU CO.
XX
XX Taniguchi T, Mizukami J;
XX
XX WPI; 1999-190635/16.
XX
XX Screening agonists/antagonists to peroxisome proliferator-activated
XX receptor - by contacting with a test cell containing a suitable reporter
XX gene.
XX
XX Example; Page 24; 34pp; Japanese.
XX
XX The invention relates to methods for identifying potential agonists or
```



XX PD 05-JUL-2001.  
XX PF 28-DEC-2000; 2000WO-US035498.  
XX PR 28-DEC-1999; 99US-0173419P.  
XX PR 27-DEC-2000; 2000US-00173419.  
XX PA (CURA-) CURAGEN CORP.  
XX PI Shimkets RA, Leach M;  
XX WPI; 2001-465210/50.  
XX CC Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,  
PT autoimmune diseases and infections.  
XX PS Claim 1; Page 3631; 4143pp; English.  
XX CC The present invention relates to oligonucleotides encoding polymorphic  
CC variants of proteins related to amyloses, amyloid proteins, angiotensin,  
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
CC histones, kinases, colony stimulating factors, complement related  
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins, G-  
CC protein coupled receptors and thioesterases. The present sequence is one  
CC such oligonucleotide, the oligonucleotides and the peptides encoded by  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate expression of the proteins listed above.  
CC Disorders that may be prevented, diagnosed and/or treated include  
CC multifactorial diseases with a genetic component, such as autoimmune  
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,  
CC leukaemia), diseases of the nervous system and an infection of pathogenic  
CC organisms  
XX SQ Sequence 50 BP; 6 A; 18 C; 16 G; 10 T; 0 U; 0 Other;  
Query Match 3.0%; Score 20.2; DB 4; Length 50;  
Best Local Similarity 88.0%; Pred. No. 7.5e+04;  
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 282 GCAGGCTCTCCGGTGCCTCCCTGCT 306  
Db 9 GCAGGCTCTCCGTGCTGCCCTGCT 33  
RESULT 12  
AAQ69713/c  
ID AAQ69713 standard; DNA; 50 BP.  
XX AC AAQ69713;  
XX DT 25-MAR-2003 (revised)  
DT 03-MAR-1995 (first entry)  
XX DE Human insulin-like growth factor II gene, target region.  
XX DNA protein-binding assay; test sequence; screening sequence; promoter;  
KW target; TATA box; Herpes Simplex Virus; HSV; origin of replication; UL9;  
KW transcription factor; TFIID; ds.  
XX OS Synthetic.  
XX PN WO9414980-A1.  
XX PD 07-JUL-1994.  
XX XX 20-DEC-1993; 93WO-US012388.  
XX PR 23-DEC-1992; 92US-00996783.  
PR 17-SEP-1993; 93US-00123936.

XX PA (GENE-) GENELABS TECHNOLOGIES INC.  
XX PI Edwards CA, Cantor CR, Andrews BM, Turin LM, Fry KE;  
XX WPI; 1994-234711/28.  
XX PT Sequence-directed DNA-binding molecules - useful in pharmaceuticals and  
PT as molecular reagents.  
XX PS Claim 28; Page 443; 587pp; English.  
XX CC A DNA protein-binding assay is provided, useful for screening libraries  
CC of synthetic or biological cpds. for their ability to bind DNA test  
CC sequences. The assay is versatile in that any number of test sequences  
CC can be tested by placing the test sequence adjacent to a defined protein-  
CC binding screening sequence. Binding of mols. to these test sequences  
CC changes the binding characteristics of the protein mol. to its cognate  
CC binding sequence. When such a mol. binds the test sequence, the  
CC equilibrium of the DNA:protein complexes is disturbed, generating changes  
CC in the concentration of free DNA probe. One application of this method is  
CC to eucaryotic general transcription factors (e.g. TFIID), where the  
CC target region is typically selected from DNA sequences adjacent to the  
CC binding site for the eucaryotic transcription factor. Numerous exemplary  
CC test sequences are given: the sequences in AAQ69251-731 and AAQ69850  
CC correspond to promoter targets (typically, TATA box-contg. sites) for  
CC human genes and the sequences in AAQ69732-849 correspond to promoter  
CC targets for viral genes. The test sequences may also be randomly  
CC generated. DNA:protein interaction may be used for screening purposes,  
CC e.g. the Herpes Simplex Virus (HSV) origin of replication and UL9 (see  
CC AAQ69851-52, AAQ69865 and AAQ69891). (Updated on 25-MAR-2003 to correct  
CC PN field.)  
XX SQ Sequence 50 BP; 6 A; 18 C; 21 G; 5 T; 0 U; 0 Other;  
Query Match 3.0%; Score 20; DB 2; Length 50;  
Best Local Similarity 65.9%; Pred. No. 8.6e+04;  
Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
QY 259 GCGGGCACACACACCTGCTCCAGCAGGCTCTCCGGTGCCTCC 302  
Db 49 GCGGGGCCACCGCCGCTCTTATAGTCGCGCCAGCCGCGGCC 6  
RESULT 13  
AAQ64175/c  
ID AAT64175 standard; DNA; 50 BP.  
XX AC AAT64175;  
XX DT 25-MAR-2003 (revised)  
DT 17-MAR-1997 (first entry)  
XX DE Human insulin-like growth factor-II gene TFIID binding site.  
XX KW Duplex DNA; target region; binding characteristic; DNA binding protein;  
KW TFIID; transcription factor; binding site; inhibition; enhance; cancer;  
KW inherited genetic disorder; ds.  
XX OS Homo sapiens.  
XX PN US5578444-A.  
XX PD 26-NOV-1996.  
XX PF 20-DEC-1993; 93US-00171389.  
XX PR 27-JUN-1991; 91US-00723618.  
PR 23-DEC-1992; 92US-00996783.  
PR 17-SEP-1993; 93US-00123936.  
XX PA (GENE-) GENELABS TECHNOLOGIES INC.  
XX XX





DR WPI; 2002-442819/47.  
 XX  
 PT Decreasing transcriptional activity of genes for treating infections or  
 PT cancer, by administration of an agent that binds to two non-overlapping  
 PT regions of the gene.  
 XX  
 PS Example 15; SEQ ID NO 453; 98pp; English.  
 XX  
 CC The invention relates to a method of decreasing transcriptional activity  
 CC in a duplex deoxyribonucleic acid (DNA) template (T1) comprising  
 CC contacting (T1) with a binding agent comprising at least one small duplex  
 CC DNA-binding molecule (T2) coupled to at least one other small duplex-  
 CC binding molecule that binds to a non-overlapping region of target  
 CC sequence (TS). The method is useful for inhibiting transcription of a  
 CC range of disease-related genes for treating infections (by viruses,  
 CC including human immunodeficiency virus, bacteria, fungi, protozoa and  
 CC parasites), cancer, cardiovascular, respiratory, gastrointestinal,  
 CC endocrine/metabolic, rheumatic/immunological, haematological,  
 CC neurological, psychiatric, dermatological, ophthalmological, musculo-  
 CC skeletal, genetic or urogenital disorders. The method provides sequence-  
 CC specific inhibition of transcription of pathological genes without  
 CC affecting transcription of cellular genes regulated by the same  
 CC transcription factor, and can be applied to regulation of any gene.  
 CC ABK82492-ABK83155 represent DNA binding molecule test sequences used in  
 CC the method of the invention  
 XX  
 SQ Sequence 50 BP; 6 A; 18 C; 21 G; 5 T; 0 U; 0 Other;  
 Query Match 3.0%; Score 20; DB 6; Length 50;  
 Best Local Similarity 65.9%; Pred. No. 8.6e+04;  
 Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
 QY 259 GCGGGCACACACCCCTCTCCAGCAGGCTCTCGGCTGCCCC 302  
 Db 49 GCGGGCGCCACGCCCGCTCTTATATCTCGCGCAGCGCGGCC 6

Search completed: February 29, 2004, 12:01:06  
 Job time : 310.274 secs



REGISTRATION NUMBER: 33,523  
REFERENCE/DOCKET NUMBER: HAZ-015CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)526-6000  
TELEFAX: (617) 526-5000  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 48 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
FRAGMENT TYPE: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 37:  
US-10-361-420-37

Query Match 3.4%; Score 23; DB 14; Length 48;  
Best Local Similarity 83.9%; Pred. No. 1.1e+04;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

304 CACCAGCGCTGAAGCTGACACCTTCAGGGAC 334  
47 CACCAGCGAGAGCCCGACCTACAGGTAC 17

RESULT 2  
US-09-736-969A-67/c  
Sequence 67, Application US/09736969A  
Patent No. US20020068302A1  
GENERAL INFORMATION:  
APPLICANT: Lu, Peter  
APPLICANT: Garman, Jonathan David  
APPLICANT: Candia III, Albert Frederick  
APPLICANT: Arbor Vita Corporation  
TITLE OF INVENTION: CLASP-4 Transmembrane Protein  
FILE REFERENCE: 020054-000411US  
CURRENT APPLICATION NUMBER: US/09/736,969A  
CURRENT FILING DATE: 2000-12-13  
PRIOR APPLICATION NUMBER: US 60/160,860  
PRIOR FILING DATE: 1999-10-21  
PRIOR APPLICATION NUMBER: US 60/162,498  
PRIOR FILING DATE: 1999-10-29  
PRIOR APPLICATION NUMBER: US 60/170,453  
PRIOR FILING DATE: 1999-12-13  
PRIOR APPLICATION NUMBER: US 60/176,195  
PRIOR FILING DATE: 2000-01-14  
PRIOR APPLICATION NUMBER: US 60/182,296  
PRIOR FILING DATE: 2000-02-14  
PRIOR APPLICATION NUMBER: US 09/547,276  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: US 60/196,267  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: US 60/196,460  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: US 60/196,527  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: US 60/196,528  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: US 09/687,837  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: US 60/240,503  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: US 60/240,508  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: US 60/240,539  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: US 60/240,543  
PRIOR FILING DATE: 2000-10-13  
NUMBER OF SEQ ID NOS: 153  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 67  
LENGTH: 50  
TYPE: DNA

ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Exon 11  
OTHER INFORMATION: ('GR2005')  
US-09-736-969A-67

Query Match 3.0%; Score 20; DB 9; Length 50;  
Best Local Similarity 72.2%; Pred. No. 7.4e+04;  
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 510 GCTCCGTCAGCGAGGCTCTGATGCAGCTGAAGC 545  
DB 44 GCTCCCGCAGCTGATGTTGCTGTCGAGCGGCATC 9

RESULT 3  
US-10-131-827-3541  
Sequence 3541, Application US/10131827  
Publication No. US20040009479A1  
GENERAL INFORMATION:  
APPLICANT: Wohlgenuth, Jay  
APPLICANT: Fry, Kirk  
APPLICANT: Woodward, Robert  
APPLICANT: Ly, Ngoc  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE  
TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES  
FILE REFERENCE: 506612000120  
CURRENT APPLICATION NUMBER: US/10/131,827  
CURRENT FILING DATE: 2002-09-06  
PRIOR APPLICATION NUMBER: US 10/006,290  
PRIOR FILING DATE: 2001-10-22  
PRIOR APPLICATION NUMBER: US 60/296,764  
PRIOR FILING DATE: 2001-06-08  
NUMBER OF SEQ ID NOS: 9090  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 3541  
LENGTH: 50  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-131-827-3541

Query Match 3.0%; Score 20; DB 15; Length 50;  
Best Local Similarity 82.1%; Pred. No. 7.4e+04;  
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 439 GGCTGGCTGGCGCATGTGCTGACTTT 466  
DB 22 GGCTGGCTCCAGATGTGTGCTTT 49

RESULT 4  
US-09-974-026-61/c  
Sequence 61, Application US/09974026  
Publication No. US20030194398A1  
GENERAL INFORMATION:  
APPLICANT: Tamburini, Paul P  
APPLICANT: Davis, Gary  
APPLICANT: Delaria, Katherine A  
APPLICANT: Christopher, Marlor W  
APPLICANT: Daniel, Muller K  
TITLE OF INVENTION: Human Bikunin  
FILE REFERENCE: 96-223-22  
CURRENT APPLICATION NUMBER: US/09/974,026  
CURRENT FILING DATE: 2001-10-10  
PRIOR APPLICATION NUMBER: US 09/144,428  
PRIOR FILING DATE: 1998-08-31  
PRIOR APPLICATION NUMBER: PCT/US97/03894  
PRIOR FILING DATE: 1997-03-10  
PRIOR APPLICATION NUMBER: US 08/725,251  
PRIOR FILING DATE: 1996-10-04  
PRIOR APPLICATION NUMBER: US 60/019,793  
PRIOR FILING DATE: 1996-06-14  
PRIOR APPLICATION NUMBER: US 60/013,106

```

; NUMBER OF SEQ ID NOS: 3090
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4141
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-827-4141

Query Match      2.9%; Score 19.6; DB 15; Length 50;
Best Local Similarity 73.5%; Pred. No. 9.5e+04;
Matches 25; Conservative 0; Mismatches 9; Indels

QY      146 AGAAGCTGCTCAAGTTTCTGTTGGTCACGACGCGTCG 179
Db      42 AGAAATTCCTGGTTAGCTCTCAGACGCGCCG 9

RESULT 7
US-09-753-436-67/c
; Sequence 67, Application US/09753436
; Patent No. US20010029293A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Vazeux, Rosemay
; TITLE OF INVENTION: ICAM-Related Materials and Methods
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Marshall, O'Tools, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/753,436
; FILING DATE:
; CLASSIFICATION:
; APPLICATION DATA:
; APPLICATION NUMBER: 09/382,289
; FILING DATE:
; APPLICATION NUMBER: US 08/487,113
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,754
; FILING DATE: 05-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/102,852
; FILING DATE: 05-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,689
; FILING DATE: 27-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Joseph A., Jr.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 33282
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856

```

INFORMATION FOR SEQ ID NO: 67:

SEQUENCE CHARACTERISTICS:  
LENGTH: 48 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA

US-09-753-436-67

Query Match 2.9%; Score 19.4; DB 9; Length 48;  
Best Local Similarity 64.4%; Pred. No. 1.1e+05;  
Matches 29; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 271 CACACTGCTCAGAGCGCCCTCCCTGCTCCCGCCACAGCCTGAA 315  
|||  
DB 45 CAGACTGCTGACATTGCACCTCGGACTGGACACCTGGCGCGCGCAA 1

RESULT 8

US-10-156-306-7723

Sequence 7723, Application US/10156306

Publication No. US20030119017A1

GENERAL INFORMATION:

APPLICANT: Ribozyme Pharmaceuticals, Inc.

APPLICANT: McSwigen, James

TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Related

FILE REFERENCE: MHE01-664-A (400/050)

CURRENT APPLICATION NUMBER: US/10/156.306

CURRENT FILING DATE: 2002-05-28

NUMBER OF SEQ ID NOS: 8013

SOFTWARE: PatentIn version 3.0

SEQ ID NO 7723

LENGTH: 48

TYPE: RNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid

US-10-156-306-7723

Query Match 2.9%; Score 19.4; DB 14; Length 48;  
Best Local Similarity 62.1%; Pred. No. 1.1e+05;  
Matches 18; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 324 CTTTCAGGAGCAGCTCCAGGAGTCTGC 352  
|||  
DB 20 CCUUCAGGACAUCGUCCGGAGCUCUGC 48

RESULT 9

US-10-163-942-67/c

Sequence 67, Application US/10163942

Publication No. US20030199423A1

GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael

Vazeux, Roseway

TITLE OF INVENTION: ICAM-Related Materials and Methods

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/163,942

FILING DATE: 05-Jun-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/753,436

FILING DATE: <Unknown>

APPLICATION NUMBER: 09/382,289

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/487,113

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/286,754

FILING DATE: 05-AUG-1994

APPLICATION NUMBER: US 08/102,852

FILING DATE: 05-AUG-1993

APPLICATION NUMBER: US 08/009,266

FILING DATE: 22-JAN-1993

APPLICATION NUMBER: US 07/894,061

FILING DATE: 05-JUN-1992

APPLICATION NUMBER: US 07/889,724

FILING DATE: 26-MAY-1992

APPLICATION NUMBER: US 07/827,689

FILING DATE: 27-JAN-1992

ATTORNEY/AGENT INFORMATION:

NAME: Williams, Joseph A., Jr.

REGISTRATION NUMBER: 38,659

REFERENCE/DOCKET NUMBER: 33282

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 474-6300

TELEFAX: (312) 474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 67:

SEQUENCE CHARACTERISTICS:

LENGTH: 48 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

SEQUENCE DESCRIPTION: SEQ ID NO: 67:

US-10-163-942-67

Query Match 2.9%; Score 19.4; DB 14; Length 48;  
Best Local Similarity 64.4%; Pred. No. 1.1e+05;  
Matches 29; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 271 CACACTGCTCAGAGCGCCCTCCCTGCTCCCGCCACAGCCTGAA 315  
|||  
DB 45 CAGACTGCTGACATTGCACCTCGGACTGGACACCTGGCGCGCGCAA 1

RESULT 10

US-09-874-547-6

Sequence 6, Application US/09874547

Patent No. US20020058269A1

GENERAL INFORMATION:

APPLICANT: No. US20020058269A1k, Steffen

APPLICANT: Kassner, Paul D.

APPLICANT: Zyomyx, Inc.

TITLE OF INVENTION: Screening of Phage Displayed Peptides

TITLE OF INVENTION: Without Clearing of the Cell Culture

FILE REFERENCE: 020144-00110US

CURRENT APPLICATION NUMBER: US/09/874,547

CURRENT FILING DATE: 2001-06-04

PRIOR APPLICATION NUMBER: 60/209,503

PRIOR FILING DATE: 2000-06-05

NUMBER OF SEQ ID NOS: 84

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 6

LENGTH: 50

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Primers for Scfv and Fab library generation (Table

OTHER INFORMATION: I) - MHC-Backs

US-09-874-547-6

Query Match 2.9%; Score 19.4; DB 9; Length 50;  
Best Local Similarity 70.3%; Pred. No. 1.1e+05;  
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 513 CCTGACGCCGAGCGTCTGATGACGCTGAAGCTTTC 549  
Db 14 CCCAGCCGCCATGCGCCAGGTGCGAGCTGAAGCAGTC 50

RESULT 11  
US-10-219-195-45/c  
; Sequence 45, Application US/10219195  
; Publication No. US20030165917A1  
; GENERAL INFORMATION:  
; APPLICANT: ULLMAN, EDWIN  
; APPLICANT: WU, MING  
; APPLICANT: LIU, YEN PING  
; TITLE OF INVENTION: ISOTHERMAL AMPLIFICATION IN NUCLEIC ACID ANALYSIS  
; FILE REFERENCE: 3817.05-1  
; CURRENT APPLICATION NUMBER: US/10/219,195  
; PRIOR FILING DATE: 2002-08-14  
; PRIOR APPLICATION NUMBER: 60/312,505  
; PRIOR FILING DATE: 2001-08-14  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 45  
; LENGTH: 46  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-219-195-45

Query Match 2.9%; Score 19.2; DB 14; Length 46;  
Best Local Similarity 67.5%; Pred. No. 1.2e+05;  
Matches 27; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 591 TCAGGAGCTCGGTACAGCGTGGCCCTGCTCCTAAAGGA 630  
Db 46 TCAAGGAGAGCTGTGCTAGCTGCCCTGGACTTCGAGCA 7

RESULT 12  
US-10-131-827-6386  
; Sequence 6386, Application US/10131827  
; Publication No. US20040009479A1  
; GENERAL INFORMATION:  
; APPLICANT: Wohlgenuth, Jay  
; APPLICANT: Fry, Kirk  
; APPLICANT: Woodward, Robert  
; APPLICANT: Ly, Ngoc  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE  
; FILE REFERENCE: 506612000120  
; CURRENT APPLICATION NUMBER: US/10/131,827  
; CURRENT FILING DATE: 2002-09-06  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 10/006,290  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/296,764  
; NUMBER OF SEQ ID NOS: 9090  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6386  
; LENGTH: 50  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-131-827-6386

Query Match 2.9%; Score 19.2; DB 15; Length 50;  
Best Local Similarity 87.5%; Pred. No. 1.2e+05;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 262 AGGATGCACACACTGCTCCAGCA 285  
Db 1 AGGAATGCACACACTGCTCCAGGA 24

RESULT 13  
US-10-131-827-6776/c  
; Sequence 6776, Application US/10131827  
; Publication No. US20040009479A1  
; GENERAL INFORMATION:  
; APPLICANT: Wohlgenuth, Jay  
; APPLICANT: Fry, Kirk  
; APPLICANT: Woodward, Robert  
; APPLICANT: Ly, Ngoc  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE  
; FILE REFERENCE: 506612000120  
; CURRENT APPLICATION NUMBER: US/10/131,827  
; CURRENT FILING DATE: 2002-09-06  
; PRIOR APPLICATION NUMBER: US 10/006,290  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/296,764  
; PRIOR FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 9090  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6776  
; LENGTH: 50  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-131-827-6776

Query Match 2.9%; Score 19.2; DB 15; Length 50;  
Best Local Similarity 87.5%; Pred. No. 1.2e+05;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 262 AGGATGCACACACTGCTCCAGCA 285  
Db 50 AGGAATGCACACACTGCTCCAGGA 27

RESULT 14  
US-09-740-668A-57/c  
; Sequence 57, Application US/09740668A  
; Patent No. US20020076700A1  
; GENERAL INFORMATION:  
; APPLICANT: Shimkets, Richard  
; TITLE OF INVENTION: No. US20020076700A1el polypeptides and nucleic acids encoding sar  
; FILE REFERENCE: 15966-537 CIP  
; CURRENT APPLICATION NUMBER: US/09/740,668A  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: PCT/US99/29584  
; PRIOR FILING DATE: 1999-12-17  
; PRIOR APPLICATION NUMBER: 09/465,512  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: 60/113,485  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/112,837  
; NUMBER OF SEQ ID NOS: 98  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 57  
; LENGTH: 36  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: 2132461 PCR reverse primer  
US-09-740-668A-57

Query Match 2.8%; Score 19; DB 9; Length 36;  
Best Local Similarity 81.5%; Pred. No. 1.3e+05;  
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 221 AGCGCCAACTGCCCGAGGAGCAGCTG 247

Db 36 AGCGTGACCTGCTGTGAGGAGAGCTG 10

RESULT 15  
US-10-429-849-12/c  
; Sequence 12, Application US/10429849  
; Publication No. US20040029157A1  
; GENERAL INFORMATION:  
; APPLICANT: TATSUKA, MASAOKI  
; APPLICANT: TERADA, YASUHIKO  
; TITLE OF INVENTION: CELL CYCLE-REGULATING PROTEINS  
; FILE REFERENCE: 050499/0102  
; CURRENT APPLICATION NUMBER: US/10/429,849  
; CURRENT FILING DATE: 2003-05-06  
; PRIOR APPLICATION NUMBER: US/09/485,534  
; PRIOR FILING DATE: 2000-02-14  
; PRIOR APPLICATION NUMBER: PCT/JP98/03641  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 235371/1997  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 38  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Primer  
US-10-429-849-12

Query Match 2.8%; Score 19; DB 12; Length 38;  
Best Local Similarity 71.4%; Pred. No. 1.4e+05;  
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 287 GCCTTCCTCTGCCCCCACCAGCCTGAAGCCTGA 321  
Db 38 GCGGCCCTCTGCTCCCTCTGCCCCCTTAGTCTGA 4

Search completed: February 29, 2004, 14:51:52  
Job time : 271.199 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 29, 2004, 11:48:29 ; Search time 68.0506 Seconds  
(without alignments)  
5488.301 Million cell updates/sec

Title: US-09-904-568-3\_COPY\_294\_966

Perfect score: 673

Sequence: 1 aatgtctgtgtggggctg.....gtgagcgcagactcaggacg 673

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 839752

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: \*

1: /cgn2\_6/prodata/2/ina/5A-COMB.seq: \*

2: /cgn2\_6/prodata/2/ina/5B-COMB.seq: \*

3: /cgn2\_6/prodata/2/ina/6A-COMB.seq: \*

4: /cgn2\_6/prodata/2/ina/6B-COMB.seq: \*

5: /cgn2\_6/prodata/2/ina/PCUS-COMB.seq: \*

6: /cgn2\_6/prodata/2/ina/backfiles.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	23	3.4	48	3	US-08-828-533-37
C 2	23	3.4	48	4	US-09-772-153-37
C 3	19.8	2.9	45	1	US-08-176-412-5
C 4	19.8	2.9	45	2	US-08-555-268A-5
C 5	19.8	2.9	45	2	US-08-495-695B-5
C 6	19.8	2.9	45	5	PCT-US94-14436-5
C 7	19.6	2.9	45	4	US-09-144-428-61
C 8	19.6	2.9	46	2	US-08-448-418-10
C 9	19.6	2.9	46	4	US-09-146-979-10
C 10	19.4	2.9	29	3	US-08-297-395-45
C 11	19.4	2.9	48	1	US-08-482-882-67
C 12	19.4	2.9	48	1	US-08-483-389-67
C 13	19.4	2.9	48	2	US-08-487-113D-67
C 14	19.4	2.9	48	2	US-08-473-503-67
C 15	19.4	2.9	48	2	US-08-483-932-67
C 16	19.4	2.9	48	2	US-08-720-420A-67
C 17	19.4	2.9	48	3	US-08-714-017-67
C 18	19.4	2.9	48	3	US-08-475-680-67
C 19	19.4	2.9	50	4	US-09-874-547-6
C 20	19	2.8	33	1	US-08-438-639-30
C 21	19	2.8	33	1	US-07-813-338A-30
C 22	19	2.8	33	3	US-08-441-971-105
C 23	19	2.8	33	3	US-08-221-653-105
C 24	19	2.8	33	3	US-08-442-144A-105
C 25	19	2.8	33	3	US-08-441-970-105
C 26	19	2.8	45	5	PCT-US94-14106-18
C 27	18.8	2.8	36	1	US-08-624-545-16

28	18.8	2.8	38	4	US-09-474-432B-1253	Sequence 1253, Ap
29	18.8	2.8	38	4	US-09-476-387-1252	Sequence 1252, Ap
30	18.8	2.8	42	1	US-08-753-054-12	Sequence 12, Appl
31	18.6	2.8	33	4	US-09-632-575-17	Sequence 17, Appl
C 32	18.6	2.8	33	4	US-09-632-575-18	Sequence 18, Appl
33	18.6	2.8	46	4	US-09-486-241-1	Sequence 1, Appli
C 34	18.6	2.8	47	4	US-09-671-317-831	Sequence 831, App
35	18.4	2.7	45	1	US-08-176-412-6	Sequence 6, Appli
36	18.4	2.7	45	2	US-08-555-268A-6	Sequence 6, Appli
37	18.4	2.7	45	2	US-08-495-695B-6	Sequence 6, Appli
38	18.4	2.7	45	5	PCT-US94-14436-6	Sequence 6, Appli
39	18.4	2.7	47	4	US-09-671-317-538	Sequence 538, App
C 40	18.4	2.7	47	4	US-09-422-978-3508	Sequence 3508, Ap
C 41	18.4	2.7	50	1	US-07-972-032-75	Sequence 75, Appl
C 42	18.2	2.7	40	2	US-08-425-884-86	Sequence 86, Appl
C 43	18.2	2.7	40	2	US-08-675-502-86	Sequence 86, Appl
C 44	18.2	2.7	40	4	US-09-245-802-86	Sequence 86, Appl
C 45	18.2	2.7	41	3	US-09-171-945-7	Sequence 7, Appli

#### ALIGNMENTS

RESULT 1  
US-08-828-533-37/c  
; Sequence 37, Application US/08828533  
; Patent No. 6180379  
; GENERAL INFORMATION:  
; APPLICANT: Ruderman, Joan V.  
; APPLICANT: Herskho, Avram  
; APPLICANT: Kirschner, Marc W.  
; APPLICANT: Townsley, Fiona  
; APPLICANT: Aristarkov, Alexander  
; APPLICANT: Eytan, Esther  
; APPLICANT: Yu, Hongtao  
; TITLE OF INVENTION: NOVEL CYCLIN-SELECTIVE UBIQUITIN CARRIER  
; TITLE OF INVENTION: POLYPEPTIDES  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HALE AND DORR LLP  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: United States of America  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/828,533  
; FILING DATE: 31-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kerner Ph.D., Ann-Louise  
; REGISTRATION NUMBER: 33,523  
; REFERENCE/DOCKET NUMBER: HAZ-015CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)526-6000  
; TELEFAX: (617) 526-5000  
; INFORMATION FOR SEQ ID NO: 37:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 48 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
; FRAGMENT TYPE: linear  
US-08-828-533-37

Query Match 3.4%; Score 23; DB 3; Length 48;  
Best Local Similarity 83.9%; Pred. No. 5.8e+03;



Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 304 CACCAGCTGAGCCTGACACCTTCAGGGAC 334  
| | | | |  
Db 47 CACCAGCCAGGAGCCCGACACCTACAGGTAC 17  
| | | | |

## RESULT 2

US-09-772-156-37/c

; Sequence 37, Application US/09772156

; Patent No. 6528633

; GENERAL INFORMATION:

; APPLICANT: Ruderman, Joan V.

; Herskho, Avram

; Kirschner, Marc W.

; Townsley, Fiona

; Aristarkov, Alexander

; Eytan, Esther

; Yu, Hongtao

; TITLE OF INVENTION: NOVEL CYCLIN-SELECTIVE UBIQUITIN CARRIER

; POLYPEPTIDES

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: HALE AND DORR LLP

; STREET: 60 State Street

; CITY: Boston

; STATE: MA

; COUNTRY: United States of America

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/772,156

; FILING DATE: 31-Mar-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/828,533

; FILING DATE: 1997-03-31

; ATTORNEY/AGENT INFORMATION:

; NAME: Kerner Ph.D., Ann-Louise

; REGISTRATION NUMBER: 33,523

; REFERENCE/DOCKET NUMBER: HA2-015CIP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 526-5000

; TELEFAX: (617) 526-5000

; INFORMATION FOR SEQ ID NO: 37:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 48 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: other nucleic acid

; FRAGMENT TYPE: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 37:

US-09-772-156-37

Query Match 3.4%; Score 23; DB 4; Length 48;

Best Local Similarity 83.9%; Pred. No. 5.8e+03;

Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 304 CACCAGCTGAGCCTGACACCTTCAGGGAC 334  
| | | | |  
Db 47 CACCAGCCAGGAGCCCGACACCTACAGGTAC 17  
| | | | |

## RESULT 3

US-08-176-412-5/c

; Sequence 5, Application US/08176412

; Patent No. 5516653

; GENERAL INFORMATION:

APPLICANT: Bard, Jonathan A.  
APPLICANT: Walker, Mary  
APPLICANT: Branchek, Theresa  
APPLICANT: Weinshank, Richard L.  
TITLE OF INVENTION: DNA ENCODING A HUMAN NEUROPEPTIDE  
TITLE OF INVENTION: V/PEPTIDE YY/PANCREATIC POLYPEPTIDE  
TITLE OF INVENTION: RECEPTOR (Y4) AND USES THEREOF  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/176,412

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 44743/JFW/TEP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 977-9550

TELEFAX: (212) 664-0525

TELEX: 422523 COOP UI

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 45 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-176-412-5

Query Match 2.9%; Score 19.8; DB 1; Length 45;

Best Local Similarity 69.2%; Pred. No. 3.3e+04;

Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 131 GACAGGAGCAGTTCAGAAAGTTGCTGAGTTTGTGTC 169  
| | | | |  
Db 44 GAGAGTTCAGATCAGATGTTGTCATTCATC 6  
| | | | |

## RESULT 4

US-08-555-268A-5/c

; Sequence 5, Application US/08555268A

; Patent No. 5958709

; GENERAL INFORMATION:

; APPLICANT: Bard, Jonathan A.

; APPLICANT: Walker, Mary

; APPLICANT: Branchek, Theresa

; APPLICANT: Weinshank, Richard L.

; TITLE OF INVENTION: DNA ENCODING A HUMAN NEUROPEPTIDE

; TITLE OF INVENTION: V/PEPTIDE YY/PANCREATIC POLYPEPTIDE

; TITLE OF INVENTION: RECEPTOR (Y4) AND USES THEREOF

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper &amp; Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:



RESULT 7  
US-09-144-428-61/c  
; Sequence 61, Application US/09144428  
; Patent No. 6583108  
; GENERAL INFORMATION:  
; APPLICANT: BAYER CORPORATION, The  
; APPLICANT: TAMBURINI, Paul P  
; APPLICANT: DAVIS, Gary  
; APPLICANT: DELARIA, Katherine A  
; APPLICANT: MARLOR, Christopher W  
; APPLICANT: MULLER, Daniel K  
; TITLE OF INVENTION: HUMAN BIKUNIN  
; NUMBER OF SEQUENCES: 71  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 S. Wacker Drive Suite 3200  
; CITY: CHICAGO  
; STATE: ILLINOIS  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/144,428  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/US97/03894  
; FILING DATE: 10-MAR-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/013,106  
; FILING DATE: 11-MAR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/019,793  
; FILING DATE: 14-JUN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/725,251  
; FILING DATE: 04-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CHAO, Mark  
; REGISTRATION NUMBER: 37,293  
; REFERENCE/DOCKET NUMBER: 96,223-II  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 913-0001  
; TELEFAX: (312) 913-0002  
; INFORMATION FOR SEQ ID NO: 61:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 45 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: DNA (genomic)  
; US-09-144-428-61  
Query Match 2.9%; Score 19.6; DB 4; Length 45;  
Best Local Similarity 73.5%; Pred. No. 3.7e+04;  
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 167 GTCAGCCTGCGAGGGGAGGACTGCGGAGCG 200  
DB 36 GTGCGCCATCTGCGAGGGCGCGGTCAGCGAGACG 3  
RESULT 8  
US-08-448-418-10  
; Sequence 10, Application US/08448418  
; Patent No. 5837242  
; GENERAL INFORMATION:  
; APPLICANT: Holliger, Kaspar-Philipp

APPLICANT: Griffiths, Andrew D  
APPLICANT: Hoogenboom, Hendricus RJM  
APPLICANT: Malmqvist, Magnus  
APPLICANT: Marks, James D  
APPLICANT: McGuinness, Brian T  
APPLICANT: Pope, Anthony R  
APPLICANT: Prospero, Terence D  
APPLICANT: Winter, Gregory P  
TITLE OF INVENTION: Multivalent and Multispecific Binding  
TITLE OF INVENTION: Proteins, Their Manufacture and Use  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall O'Toole Gerstein Murray and Borun  
STREET: 6300 Sears Tower 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/448,418  
FILING DATE: 14-MAY-1996  
CLASSIFICATION: C12N 15/62, 15/70, C07K 1/00  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB93/02492  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9225453.1  
FILING DATE: 04-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9300816.7  
FILING DATE: 16-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 93303614.7  
FILING DATE: 10-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9319969.3  
FILING DATE: 22-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: David W. Clough  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 28111/32651  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 46 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA primer  
US-08-448-418-10  
Query Match 2.9%; Score 19.6; DB 2; Length 46;  
Best Local Similarity 66.7%; Pred. No. 3.7e+04;  
Matches 28; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
QY 510 GTCCTCTGAGCGCGGCTCTGTGTCAGCTGAGCTTCAG 551  
DB 4 GCGGCCAGCGCGCCATGCGCCAGGTGTCAGCTGAGGAGTCAG 45  
RESULT 9  
US-09-146-979-10  
; Sequence 10, Application US/09146979  
; Patent No. 6492123  
; GENERAL INFORMATION:  
; APPLICANT: Holliger, Kaspar-Philipp  
; APPLICANT: Griffiths, Andrew D  
; APPLICANT: Hoogenboom, Hendricus RJM  
; APPLICANT: Malmqvist, Magnus

APPLICANT: Marks, James D  
APPLICANT: McGuinness, Brian T  
APPLICANT: Pope, Anthony R  
APPLICANT: Prospero, Terence D  
APPLICANT: Winter, Gregory P  
TITLE OF INVENTION: Multivalent and Multispecific Binding  
TITLE OF INVENTION: Proteins, Their Manufacture and Use  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall O'Toole Gerstein Murray and Borun  
STREET: 6300 Sears Tower 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/146,979  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/448,418  
FILING DATE: 14-MAY-1996  
APPLICATION NUMBER: PCT/GB93/02492  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9225453.1  
FILING DATE: 04-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9300816.7  
FILING DATE: 16-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 93303614.7  
FILING DATE: 10-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9319969.3  
FILING DATE: 22-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: David W. Clough  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 28111/32651  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 46 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA primer  
US-09-146-979-10  
Query Match 2.9%; Score 19.6; DB 4; Length 46;  
Best Local Similarity 66.7%; Pred. No. 3.7e+04;  
Matches 28; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
QY 510 GTCCTCGCCGCGCGCTCTGATGAGCTGAGCTTTTCAG 551  
DB 4 GCGGCCAGCGCGCGCATGGCCGAGTGCAGCTGAAGGAGTCAG 45  
RESULT 10  
US-08-297-395-45/c  
Sequence 45, Application US/08297395A  
Patent No. 6039947  
GENERAL INFORMATION:  
APPLICANT: Howard L. Weiner  
APPLICANT: David A. Hafner  
TITLE OF INVENTION: EPITOPES DERIVED FROM IMMUNODOMINANT  
TITLE OF INVENTION: EPITOPES OF MYELIN BASIC PROTEIN  
FILE REFERENCE: 1010/05723US3

CURRENT APPLICATION NUMBER: US/08/297,395A  
CURRENT FILING DATE: 1994-08-11  
EARLIER APPLICATION NUMBER: 08/059,189  
EARLIER FILING DATE: 1993-05-06  
EARLIER APPLICATION NUMBER: 07/502,559  
EARLIER FILING DATE: 1990-03-30  
EARLIER APPLICATION NUMBER: PCT/US88/02139  
EARLIER FILING DATE: 1988-06-24  
EARLIER APPLICATION NUMBER: 07/065,734  
EARLIER FILING DATE: 1987-06-24  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 45  
LENGTH: 29  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-08-297-395-45  
Query Match 2.9%; Score 19.4; DB 3; Length 29;  
Best Local Similarity 79.3%; Pred. No. 3.7e+04;  
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 422 GTGCCCCAGCAGCAGCGGCGCTGCTGCC 450  
DB 29 GTGTCTTACCAGCAAGGCTCTGTCTGCC 1  
RESULT 11  
US-08-482-882-67/c  
Sequence 67, Application US/08482882  
Patent No. 5773218  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Vazeux, Rosemay  
TITLE OF INVENTION: ICAM-Related Materials and Methods  
NUMBER OF SEQUENCES: 116  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,882  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,754  
FILING DATE:  
APPLICATION NUMBER: US 08/102,852  
FILING DATE: 05-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/009,266  
FILING DATE: 22-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/894,061  
FILING DATE: 05-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/889,724  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,689  
FILING DATE: 27-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5773218and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 32178

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 48 base pairs  
TYPE: nucleic acid  
STRADEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-483-882-67

Query Match 2.9%; Score 19.4; DB 1; Length 48;  
Best Local Similarity 64.4%; Pred. No. 4.2e+04;  
Matches 29; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 271 CACACTGCTCCAGCAGGCGCTCGTCTGCCCGCCACACCTGAA 315  
Db 45 CAGACTGCTGCACTTGCACCTCGGACTGGACACCTGGCGCGCGCA 1

RESULT 12  
US-08-483-389-67/c  
Sequence 67, Application US/08483389  
Patent No. 5811517  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Vazeux, Rosemay  
TITLE OF INVENTION: ICAM-RELATED PROTEIN  
NUMBER OF SEQUENCES: 118  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive/6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION NUMBER: US 08/483,389  
APPLICATION NUMBER: US 08/102,852  
FILING DATE: 05-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/009,266  
FILING DATE: 22-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/889,724  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/894,061  
FILING DATE: 05-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/889,724  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,689  
FILING DATE: 27-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Suh, Young J.  
REGISTRATION NUMBER: P-41,337  
REFERENCE/DOCKET NUMBER: 27866/32760  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: (312) 474-6600  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 48 base pairs

TYPE: nucleic acid  
STRADEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-483-389-67

Query Match 2.9%; Score 19.4; DB 1; Length 48;  
Best Local Similarity 64.4%; Pred. No. 4.2e+04;  
Matches 29; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 271 CACACTGCTCCAGCAGGCGCTCGTCTGCCCGCCACACCTGAA 315  
Db 45 CAGACTGCTGCACTTGCACCTCGGACTGGACACCTGGCGCGCGCA 1

RESULT 13  
US-08-487-113D-67/c  
Sequence 67, Application US/08487113D  
Patent No. 5837822  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Vazeux, Rosemay  
TITLE OF INVENTION: ICAM-Related Materials and Methods  
NUMBER OF SEQUENCES: 120  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-8402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: US/08/487,113D  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,754  
FILING DATE: 05-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/102,852  
FILING DATE: 05-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/009,266  
FILING DATE: 22-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/894,061  
FILING DATE: 05-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/889,724  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,689  
FILING DATE: 27-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5837822 and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 32744  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 48 base pairs  
TYPE: nucleic acid  
STRADEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA

US-08-487-113D-67

Query Match 2.9%; Score 19.4; DB 2; Length 48;  
Best Local Similarity 64.4%; Pred. No. 4.2e+04;  
Matches 29; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 271 CACACTGCTCCAGCAGCCCTCGCTCGCCCCCACCAGCCTGAA 315  
DB 45 CAGACTGCTGAGTTGCACCTCGGACTGGACACCTGGCGCGCAA 1

RESULT 14

US-08-473-503-67/c  
; Sequence 67, Application US/08473503  
; Patent No. 5869262  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Vazeux, Rosemay  
; TITLE OF INVENTION: ICAM-Related Materials and Methods  
; NUMBER OF SEQUENCES: 116  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/473,503  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/286,754  
; FILING DATE: 05-AUG-1994  
; APPLICATION NUMBER: US 08/102,852  
; FILING DATE: 05-AUG-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/009,266  
; FILING DATE: 22-JAN-1993  
; APPLICATION NUMBER: US 07/894,061  
; FILING DATE: 05-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/889,724  
; FILING DATE: 26-MAY-1992  
; APPLICATION NUMBER: US 07/827,689  
; FILING DATE: 27-JAN-1992  
; APPLICATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 32178  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 67:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 48 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-08-473-503-67

Query Match 2.9%; Score 19.4; DB 2; Length 48;  
Best Local Similarity 64.4%; Pred. No. 4.2e+04;  
Matches 29; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 271 CACACTGCTCCAGCAGCCCTCGCTCGCCCCCACCAGCCTGAA 315  
DB 45 CAGACTGCTGAGTTGCACCTCGGACTGGACACCTGGCGCGCAA 1

RESULT 15

US-08-483-932-67/c  
; Sequence 67, Application US/08483932  
; Patent No. 5880268  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Vazeux, Rosemay  
; TITLE OF INVENTION: ICAM-Related Materials and Methods  
; NUMBER OF SEQUENCES: 116  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/483,932  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/286,754  
; FILING DATE: 05-AUG-1994  
; APPLICATION NUMBER: US 08/102,852  
; FILING DATE: 05-AUG-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/009,266  
; FILING DATE: 22-JAN-1993  
; APPLICATION NUMBER: US 07/894,061  
; FILING DATE: 05-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/889,724  
; FILING DATE: 26-MAY-1992  
; APPLICATION NUMBER: US 07/827,689  
; FILING DATE: 27-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5880268and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 32178  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 67:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 48 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-08-483-932-67

Query Match 2.9%; Score 19.4; DB 2; Length 48;  
Best Local Similarity 64.4%; Pred. No. 4.2e+04;  
Matches 29; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 271 CACACTGCTCCAGCAGCCCTCGCTCGCCCCCACCAGCCTGAA 315  
DB 45 CAGACTGCTGAGTTGCACCTCGGACTGGACACCTGGCGCGCAA 1

Mon Mar 1 10:00:10 2004

us-09-904-568-3\_copy\_294\_966.szlm50.rni

Page 8

Search completed: February 29, 2004, 14:42:41  
Job time : 70.0506 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 29, 2004, 10:34:43 ; Search time 2034.51 Seconds  
(without alignments)  
9878.168 Million cell updates/sec

Title: US-09-904-568-3\_COPY\_294\_966

Perfect score: 673

Sequence: 1 aatgtctgtg9gggctg.....gtgagcgagactgcaggac 673

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 138346

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estm:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_nam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
C 1	21.8	3.2	50 9	AU102670
C 2	20.8	3.1	50 29	AL946896
C 3	20.6	3.1	45 12	BG772081
C 4	20	3.0	46 14	H45385

5	19.8	2.9	45	14	T63738	T63738 yc23e03.r1
6	19.8	2.9	46	9	AI153463	AI153463 uc53e11.r
C 7	19.8	2.9	50	9	AU102937	AU102937 AU102937
C 8	19.8	2.9	50	9	AU103175	AU103175 AU103175
C 9	19.8	2.9	50	9	AU106019	AU106019 AU106019
C 10	19.6	2.9	50	9	AL651763	AL651763 AL651763
C 11	19.4	2.9	40	12	BI160469	BI160469 602864532
12	19.2	2.9	43	29	CG714548	CG714548 1119037B1
13	19.2	2.9	44	29	CG708599	CG708599 1119010A0
14	19.2	2.9	44	29	CG708605	CG708605 1119010A0
15	19.2	2.9	44	29	CG708611	CG708611 1119010A0
16	19.2	2.9	44	29	CG708616	CG708616 1119010A0
17	19.2	2.9	44	29	CG708624	CG708624 1119010A1
18	19.2	2.9	44	29	CG708627	CG708627 1119010A1
19	19.2	2.9	44	29	CG708670	CG708670 1119010C0
20	19.2	2.9	44	29	CG708679	CG708679 1119010C0
21	19.2	2.9	44	29	CG708685	CG708685 1119010C0
22	19.2	2.9	44	29	CG708716	CG708716 1119010D0
23	19.2	2.9	44	29	CG708728	CG708728 1119010D1
24	19.2	2.9	44	29	CG708745	CG708745 1119010E0
25	19.2	2.9	44	29	CG708750	CG708750 1119010E0
26	19.2	2.9	44	29	CG708753	CG708753 1119010E1
27	19.2	2.9	44	29	CG708770	CG708770 1119010F0
28	19.2	2.9	44	29	CG708785	CG708785 1119010F0
29	19.2	2.9	44	29	CG708791	CG708791 1119010F1
30	19.2	2.9	44	29	CG708817	CG708817 1119010G0
31	19.2	2.9	44	29	CG708833	CG708833 1119010H0
32	19.2	2.9	44	29	CG708836	CG708836 1119010H0
33	19.2	2.9	44	29	CG708839	CG708839 1119010H0
34	19.2	2.9	44	29	CG708859	CG708859 1119010H1
35	19.2	2.9	44	29	CG708865	CG708865 1119011A0
36	19.2	2.9	44	29	CG708904	CG708904 1119011B0
37	19.2	2.9	44	29	CG708937	CG708937 1119011C0
38	19.2	2.9	44	29	CG708940	CG708940 1119011C0
39	19.2	2.9	44	29	CG708943	CG708943 1119011C0
40	19.2	2.9	44	29	CG708980	CG708980 1119011D0
41	19.2	2.9	44	29	CG708989	CG708989 1119011D1
42	19.2	2.9	44	29	CG708999	CG708999 1119011E0
43	19.2	2.9	44	29	CG709002	CG709002 1119011E0
44	19.2	2.9	44	29	CG709025	CG709025 1119011E1
45	19.2	2.9	44	29	CG709028	CG709028 1119011F0

ALIGNMENTS

RESULT 1  
AU102670/c  
LOCUS AU102670 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone  
DEFINITION CAS07213, mRNA sequence.  
ACCESSION AU102670  
VERSION AU102670.1 GI:13552191  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 50)  
AUTHORS Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J.,  
Kata, H., Oka, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K.,  
Sakaki, Y., Nakamura, Y., Suyama, A. and Sugano, S.  
Diverse transcriptional initiation revealed by fine, large-scale  
mapping of mRNA start sites  
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)  
MEDLINE 21270072  
PUBMED 11375929  
COMMENT Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: yusuzuki@ims.u-tokyo.ac.jp  
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and





```

DEFINITION YN99cl2.r1 Soares adult brain N2b5HB55Y Homo sapiens cDNA clone
IMAGE:176566 5', similar to SP:TF52_MOUSE P10712 TRANSCRIPTION
FACTOR S-II ;, mRNA sequence.
H45385
KEYWORDS H45385.1 GI:921437
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 46)
REFERENCE Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
AUTHORS Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: M13RPI
High quality sequence stop: 1.
Location/Qualifiers
1..46
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3838762"
/db_xref="taxon:9606"
/clone="IMAGE:176566"
/sex="Male"
/dev_stage="55-year old"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares adult brain N2b5HB55Y"
/notes="Organ: brain; Vector: pVT3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTGTGTTTGTGTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pVT3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 53. Library constructed by Bento
Soares and M.Fatima Bonaldo. The adult brain RNA was
provided by Dr. Donald H. Gilden. Tissue was acquired
17-18 hours after death which occurred in consequence of a
ruptured aortic aneurysm. RNA was prepared from a pool of
tissues representing the following areas of the brain:
frontal, parietal, temporal and occipital cortex from the
left and right hemispheres, subcortical white matter,
basal ganglia, thalamus, cerebellum, midbrain, pons and
medulla."
ORIGIN
Query Match 3.0%; Score 20; DB 14; Length 46;
Best Local Similarity 72.2%; Pred. No. 2.3e+06;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 351 GCATCCCCAAGACCTGGTGGGACTTCGCCAGCG 386
DB 41 GCTTCTCCAGACTTGATGAGAGACTTCGCCAGTG 6

```

---

```

RESULT 5
LOCUS T63738
DEFINITION YC23e03.r1 Stragatene lung (#937210) Homo sapiens cDNA clone
IMAGE:81532 5', similar to gb:X55715 40S RIBOSOMAL PROTEIN S3
(HUMAN);, mRNA sequence.
T63738 45 bp mRNA linear EST 17-FEB-1995
T63738.1 GI:667603
ACCESSION T63738
VERSION T63738.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 45)
REFERENCE Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
AUTHORS Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevasakis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
8889549
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 55
High quality sequence stops: 37
Source: IMAGE Consortium, LLNL This
clone is available royalty-free through LLNL; contact the IMAGE
Consortium (info@image.llnl.gov) for further information.
Insert length: 55 Std Error: 0.00
Seq primer: M13RPI
High quality sequence stop: 37.
Location/Qualifiers
1..45
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:485149"
/db_xref="taxon:9606"
/clone="IMAGE:81532"
/sex="male"
/dev_stage="72 years"
/lab_host="SOLR cells (kanamycin resistant)"
/clone_lib="Stragatene lung (#937210)"
/notes="Organ: lung; Vector: pBluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. normal lung. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGACGAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
ORIGIN
Query Match 2.9%; Score 19.8; DB 14; Length 45;
Best Local Similarity 77.4%; Pred. No. 2.5e+06;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 628 GGAGATCGCAGATCTCGAGAAGAGGTGTGTGAG 658
DB 1 GAAGATGGCTACTCTGGAGTTGAGTCCGAG 31

```

---

```

RESULT 6
LOCUS A1153463
DEFINITION uc53e11.r1 Soares thymus 2NDMT Mus musculus cDNA clone
IMAGE:1429388 5', similar to SW:PHLX_RABIT Q05017 PHOSPHOLIPASE
ADRB-B PRECURSOR ;, mRNA sequence.
A1153463 46 bp mRNA linear EST 30-SEP-1998
A1153463.1 GI:3681932
ACCESSION A1153463
VERSION A1153463
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 46)
REFERENCE Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
AUTHORS Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: M13RPI
High quality sequence stop: 1.
Location/Qualifiers
1..46
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3838762"
/db_xref="taxon:9606"
/clone="IMAGE:176566"
/sex="Male"
/dev_stage="55-year old"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares adult brain N2b5HB55Y"
/notes="Organ: brain; Vector: pVT3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTGTGTTTGTGTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pVT3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 53. Library constructed by Bento
Soares and M.Fatima Bonaldo. The adult brain RNA was
provided by Dr. Donald H. Gilden. Tissue was acquired
17-18 hours after death which occurred in consequence of a
ruptured aortic aneurysm. RNA was prepared from a pool of
tissues representing the following areas of the brain:
frontal, parietal, temporal and occipital cortex from the
left and right hemispheres, subcortical white matter,
basal ganglia, thalamus, cerebellum, midbrain, pons and
medulla."
ORIGIN
Query Match 3.0%; Score 20; DB 14; Length 46;
Best Local Similarity 72.2%; Pred. No. 2.3e+06;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 351 GCATCCCCAAGACCTGGTGGGACTTCGCCAGCG 386
DB 41 GCTTCTCCAGACTTGATGAGAGACTTCGCCAGTG 6

```



Query Match 2.9%; Score 19.8; DB 9; Length 50;  
 Best Local Similarity 63.8%; Pred. No. 2.7e+06;  
 Matches 30; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 268 GCACACATGCTCCAGCAGCCCTCGCTGCGCCGCCACACAGCTGCA 314  
 |||||  
 Db 50 GCACGCCCGCTCTGCAGCCCACTCTGTGGCGCCCTTAACCTGCA 4

RESULT 9  
 AUI06019/c  
 LOCUS AUI06019 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone  
 DEFINITION KAI01053, mRNA sequence.  
 ACCESSION AUI06019  
 VERSION AUI06019.1 GI:113555540  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 50)  
 AUTHORS Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J.,  
 Hata, H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K.,  
 Sakaki, Y., Nakamura, Y., Sugama, A. and Sugano, S.  
 TITLE Diverse transcriptional initiation revealed by fine, large-scale  
 mapping of mRNA start sites  
 JOURNAL EMBO Rep. 2 (5), 388-393 (2001)  
 MEDLINE 21270072  
 PUBMED 11375929  
 COMMENT Contact: Yutaka Suzuki  
 Department of Virology  
 Institute of Medical Science, University of Tokyo  
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
 Email: yusuk@iims.u-tokyo.ac.jp  
 Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and  
 Sugano, S. Construction and characterization of a full  
 length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),  
 149-156 (1997).

FEATURES  
 source  
 1..50  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="KAI01053"  
 /clone\_lib="Sugano Homo sapiens cDNA library"

ORIGIN  
 Query Match 2.9%; Score 19.8; DB 9; Length 50;  
 Best Local Similarity 63.8%; Pred. No. 2.7e+06;  
 Matches 30; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 236 GAGCAGCAGCTGGTGGCTCGCTGGCAGGAGCATGCACACAGCTGCCA 282  
 |||||  
 Db 50 GAGAGCAGCAGCGGGCAGGAGGAGCGGCGGCGAATACACAGCTGCTTA 4

RESULT 10  
 AL651763/c  
 LOCUS AL651763 XGC-gastrula Silurana tropicalis cDNA clone TGA036n06 5',  
 DEFINITION mRNA sequence.  
 ACCESSION AL651763  
 VERSION AL651763.1 GI:17661998  
 KEYWORDS EST.  
 SOURCE Silurana tropicalis (western clawed frog)  
 ORGANISM Silurana tropicalis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 Xenopodinae; Silurana.  
 REFERENCE 1 (bases 1 to 50)  
 AUTHORS Croning, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.

TITLE Sanger Xenopus tropicalis EST project 2001 (11\_2003)  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Huckle E  
 Sanger Institute  
 Hinxton, Cambridgeshire, CB10 1SA, UK  
 Email: tropesanger.ac.uk  
 This sequence is from a Xenopus Gene Collection (XGC) library  
 constructed by Aaron M. Zorn.  
 cDNA was oligo dt primed from sug of poly A+ RNA from stages 10-13  
 gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with  
 EcoRI at the 5' end and NotI at the 3' end.  
 Vector: pCS107; Site 1: EcoRI; Site 2: NotI  
 Host: Escherichia coli XL1-blue  
 Sanger Xenopus tropicalis EST project 2001  
 TROPICALIS\_SEQUENCE\_ID: TGA036n06.plkSF6  
 Sequencing primer: SP6.  
 Location/Qualifiers  
 1..50  
 /organism="Silurana tropicalis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8364"  
 /clone="TGA036n06"  
 /dev\_stage="gastrula (stages 10.5-12 mixed)"  
 /lab\_host="Escherichia coli XL1-blue"  
 /clone\_lib="XGC-gastrula"  
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA  
 was oligo dt primed from sug of poly A+ RNA from stages  
 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated  
 into pCS107 with EcoRI at the 5' end and NotI at the 3'  
 end."

ORIGIN  
 Query Match 2.9%; Score 19.6; DB 9; Length 50;  
 Best Local Similarity 62.0%; Pred. No. 2.9e+06;  
 Matches 31; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 201 GTGTGCAGCGTCTTGGGTCAGGCCCACTGCGGAGGAGCAGCTGGGT 250  
 |||||  
 Db 50 GTGTGCCCCGTATGGGGCATCTGTGCACCTGCCCGCCCATCAGGTT 1

RESULT 11  
 B1160469/c  
 LOCUS B1160469 42 Homo sapiens cDNA clone IMAGE:5018743 5',  
 DEFINITION mRNA sequence.  
 ACCESSION B1160469  
 VERSION B1160469.1 GI:14620470  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 40)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM1832 row: p column: 08  
 High quality sequence stop: 40.  
 Location/Qualifiers  
 1..40  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"

/clone="IMAGE:5018743"  
 /tissue\_type="epithelioid carcinoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_42"  
 /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;  
 Site 2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Size-selected 500bp  
 for average insert size 1.8kb. Library constructed by Ling  
 Hong in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 2.9%; Score 19.4; DB 12; Length 40;  
 Best Local Similarity 70.3%; Pred. No. 2.9e+06;  
 Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 346 GCTCTGCTATCCCAAGACCTGCTGGGACTTGGCC 382

Db 37 GCTCGCAGCCCGGACCCGCTCGGGGATCTCGCC 1

## RESULT 12

LOCUS CG714548 43 bp DNA linear GSS 20-OCT-2003  
 DEFINITION 1119037B11.1EL.x1 1119 - RescueMu Grid AA Zea mays genomic, genomic  
 survey sequence.

ACCESSION CG714548

VERSION CG714548.1 GI:37741009

KEYWORDS GSS.

SOURCE Zea mays

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 43)

AUTHORS Walbot V.

TITLE Maize genomic sequences found using engineered RescueMu transposon

JOURNAL Unpublished (2001)

COMMENT Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Very probable ligation site of ends cut by single endonuclease.

Reverse complemented post-ligation sequence from source sequence.

Plate: 119037 row: 32

Class: transposon-tagged.

Location/Qualifiers

## FEATURES

source

1..43

/organism="Zea mays"

/mol\_type="genomic DNA"

/cultivar="mixed background W23/A188/B73/K55"

/db\_xref="taxon:4577"

/tissue\_type="leaf"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_lib="1119 - RescueMu Grid AA"

/note="Organ: leaf; Vector: RescueMu (engineered from

pBluescript backbone); Site 1: BamHI; Site 2: BglII;

RescueMu is a 4.9 kb, modified maize Mu transposon

designed to allow plasmid rescue from total genomic DNA.

Mu elements insert preferentially into transcription

units. For more information on RescueMu, go to the web

site 'www.zmdb.iastate.edu' and follow the links for

'RescueMu.' Grid AA was grown at UC San Diego in 2002. DNA

was extracted from leaf strips, double digested using

BamHI and BglII, and ligated to form circular plasmids.

DH10B cells were transformed and then screened on LB

plates with ampicillin."

## ORIGIN

Query Match 2.9%; Score 19.2; DB 29; Length 43;  
 Best Local Similarity 67.5%; Pred. No. 3.3e+06;  
 Matches 27; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 229 CTTGCCGAGGAGCAGCTGGTGGCCCTGCTGGCAGGCATG 268

Db 1 CTTGCCGAGGAGCAGCAGCATCTCATCTCTGAGCGACAGG 40

## RESULT 13

CG708599

LOCUS

DEFINITION

1119010A02.1EL.x1 1119 - RescueMu Grid AA Zea mays genomic, genomic

survey sequence.

ACCESSION CG708599

VERSION CG708599.1 GI:37734505

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 44)

AUTHORS Walbot V.

TITLE Maize genomic sequences found using engineered RescueMu transposon

JOURNAL Unpublished (2001)

COMMENT Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Very probable ligation site of ends cut by single endonuclease.

Reverse complemented post-ligation sequence from source sequence.

Plate: 1119010 row: 33

Class: transposon-tagged.

Location/Qualifiers

1..44

/organism="Zea mays"

/mol\_type="genomic DNA"

/cultivar="mixed background W23/A188/B73/K55"

/db\_xref="taxon:4577"

/tissue\_type="leaf"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_lib="1119 - RescueMu Grid AA"

/note="Organ: leaf; Vector: RescueMu (engineered from

pBluescript backbone); Site 1: BamHI; Site 2: BglII;

RescueMu is a 4.9 kb, modified maize Mu transposon

designed to allow plasmid rescue from total genomic DNA.

Mu elements insert preferentially into transcription

units. For more information on RescueMu, go to the web

site 'www.zmdb.iastate.edu' and follow the links for

'RescueMu.' Grid AA was grown at UC San Diego in 2002. DNA

was extracted from leaf strips, double digested using

BamHI and BglII, and ligated to form circular plasmids.

DH10B cells were transformed and then screened on LB

plates with ampicillin."

## ORIGIN

Query Match 2.9%; Score 19.2; DB 29; Length 44;  
 Best Local Similarity 67.5%; Pred. No. 3.3e+06;  
 Matches 27; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 229 CTTGCCGAGGAGCAGCTGGTGGCCCTGCTGGCAGGCATG 268

Db 2 CTTGCCGAGGAGCAGCAGCATCTCATCTCTGAGCGACAGG 41



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 29, 2004, 10:05:33 : Search time 2729.03 Seconds  
(without alignments)  
10688.739 Million cell updates/sec

Title: US-09-904-568-3\_COPY\_294\_966

Perfect score: 673

Sequence: 1 aatgtctgctgtggtgggctg.....gtgagcgagactgcaggac 673

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 1603530

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*\*

1: gb\_ba:\*\*

2: gb\_htg:\*\*

3: gb\_in:\*\*

4: gb\_on:\*\*

5: gb\_ov:\*\*

6: gb\_pat:\*\*

7: gb\_ph:\*\*

8: gb\_pl:\*\*

9: gb\_pr:\*\*

10: gb\_ro:\*\*

11: gb\_sts:\*\*

12: gb\_sy:\*\*

13: gb\_un:\*\*

14: gb\_vi:\*\*

15: em\_ba:\*\*

16: em\_fun:\*\*

17: em\_hum:\*\*

18: em\_in:\*\*

19: em\_mu:\*\*

20: em\_mu:\*\*

21: em\_or:\*\*

22: em\_ov:\*\*

23: em\_pat:\*\*

24: em\_ph:\*\*

25: em\_pl:\*\*

26: em\_ro:\*\*

27: em\_sts:\*\*

28: em\_un:\*\*

29: em\_vi:\*\*

30: em\_htg\_hum:\*\*

31: em\_htg\_inv:\*\*

32: em\_htg\_other:\*\*

33: em\_htg\_mus:\*\*

34: em\_htg\_pin:\*\*

35: em\_htg\_rod:\*\*

36: em\_htg\_man:\*\*

37: em\_htg\_vrt:\*\*

38: em\_sy:\*\*

39: em\_htgo\_hum:\*\*

40: em\_htgo\_mus:\*\*

41: em\_htgo\_other:\*\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	23	3.4	48	6	AR127012	AR127012 Sequence
C 2	23	3.4	48	6	AR285578	AR285578 Sequence
C 3	21	3.1	45	6	A76101	A76101 Sequence 40
C 4	20.6	3.1	40	6	AX517113	AX517113 Sequence
C 5	20.6	3.1	40	6	AX519642	AX519642 Sequence
C 6	20.2	3.0	42	6	EA1550	EA1550 Method for
C 7	20.2	3.0	42	6	BD013091	BD013091 Method for
C 8	20	3.0	50	5	CHKGDCG	D90240 Gallus gall
C 9	20	3.0	50	6	AX173088	AX173088 Sequence
C 10	19.8	2.9	45	6	AR075912	AR075912 Sequence
C 11	19.8	2.9	45	6	AR083198	AR083198 Sequence
C 12	19.8	2.9	45	6	BD247107	BD247107 Improved
C 13	19.8	2.9	45	6	I20781	I20781 Sequence 5
C 14	19.8	2.9	48	6	BD237452	BD237452 Single nu
C 15	19.8	2.9	48	6	BD270516	BD270516 Method. 7
C 16	19.8	2.9	48	6	AX026686	AX026686 Sequence
C 17	19.8	2.9	50	9	AF057510	AF057510 Homo sapi
C 18	19.6	2.9	41	6	AX513948	AX513948 Sequence
C 19	19.6	2.9	41	6	AX513287	AX513287 Sequence
C 20	19.6	2.9	45	6	BD247648	BD247648 A method
C 21	19.6	2.9	45	6	AR345117	AR345117 Sequence
C 22	19.6	2.9	46	6	A38923	A38923 Sequence 10
C 23	19.6	2.9	46	6	AR054319	AR054319 Sequence
C 24	19.6	2.9	46	6	AR265069	AR265069 Sequence
C 25	19.6	2.9	48	6	AX221581	AX221581 Sequence
C 26	19.4	2.9	31	6	E26067	E26067 Oligonucleo
C 27	19.4	2.9	48	6	AR013871	AR013871 Sequence
C 28	19.4	2.9	48	6	AR033825	AR033825 Sequence
C 29	19.4	2.9	48	6	AR042485	AR042485 Sequence
C 30	19.4	2.9	48	6	AR058365	AR058365 Sequence
C 31	19.4	2.9	48	6	AR088191	AR088191 Sequence
C 32	19.4	2.9	48	6	BD237453	BD237453 Single nu
C 33	19.4	2.9	48	6	BD270517	BD270517 Method. 7
C 34	19.4	2.9	48	6	AX026687	AX026687 Sequence
C 35	19.4	2.9	48	6	AX583009	AX583009 Sequence
C 36	19.4	2.9	50	6	AX157594	AX157594 Sequence
C 37	19.4	2.9	50	6	AX351171	AX351171 Sequence
C 38	19.2	2.9	41	6	AX516097	AX516097 Sequence
C 39	19.2	2.9	41	6	AX517503	AX517503 Sequence
C 40	19.2	2.9	45	6	AX840266	AX840266 Sequence
C 41	19	2.8	33	6	AR004376	AR004376 Sequence
C 42	19	2.8	33	6	AR097167	AR097167 Sequence
C 43	19	2.8	33	6	AR130665	AR130665 Sequence
C 44	19	2.8	33	6	AR172014	AR172014 Sequence
C 45	19	2.8	33	6	I82851	I82851 Sequence 30

ALIGNMENTS

RESULT 1  
AR127012/c  
LOCUS AR127012 48 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 37 from patent US 6180379.  
ACCESSION AR127012  
VERSION AR127012.1 GI:14113605  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 48)  
AUTHORS Ruderman,J.V., Hershko,A., Kirschner,M.W., Townsley,F.,  
Aristarkov,A., Eytan,E. and Yu,H.  
TITLE Cyclin-selective ubiquitin carrier polypeptides  
JOURNAL Patent: US 6180379-A 37 30-JAN-2001;

```
FEATURES
  source
    Location/Qualifiers
    1..48
    /organism="unknown"
    /mol_type="unassigned DNA"
ORIGIN
  Query Match
  Best Local Similarity 3.4%; Score 23; DB 6; Length 48;
  Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 304 CACAGCCTGAGCTGACACCTTCAGGGAC 334
Db 47 CACAGCCAGGAGCCCGACCTACAGGTAC 17
RESULT 2
LOCUS AR285578/c
DEFINITION Sequence 37 from patent US 6528633.
ACCESSION AR285578
VERSION AR285578.1 GI:29723140
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 48)
  Ruderman,J.V., Herskho,A., Kirschner,M.W., Townsley,F.,
  Aristarkov,A., Eytan,E. and Yu.H.
  Cyclin-selective ubiquitin carrier polypeptides
  Patent: US 6528633-A 37 04-MAR-2003;
  Location/Qualifiers
  1..48
  /organism="unknown"
  /mol_type="unassigned DNA"
ORIGIN
  Query Match
  Best Local Similarity 3.4%; Score 23; DB 6; Length 48;
  Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 304 CACAGCCTGAGCTGACACCTTCAGGGAC 334
Db 47 CACAGCCAGGAGCCCGACCTACAGGTAC 17
RESULT 3
LOCUS A76101/c
DEFINITION Sequence 40 from Patent WO9320210.
ACCESSION A76101
VERSION A76101.1 GI:6088242
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE
  1 (bases 1 to 45)
  Taylor,G. and Stott,E.J.
  ANTIBODIES FOR TREATMENT AND PREVENTION OF RESPIRATORY SYNCYTIAL
  VIRUS INFECTION
  Patent: WO 9320210-A 40 14-OCT-1993;
  SCOTGEN LTD (GB); TAYLOR GERALDINE (GB)
  Location/Qualifiers
  1..45
  /organism="unidentified"
  /mol_type="unassigned DNA"
  /db_xref="taxon:32644"
ORIGIN
  Query Match
  Best Local Similarity 3.1%; Score 21; DB 6; Length 45;
  Matches 30; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 42 CTGGTGATAGTACAGTGGCCGAGTGAGTCTTTCTTGGGGGCCAGC 86
Db 45 CTGCTGATCTACAGAGTTCCTCAACCGAATTTCTGTGTGCCAAGC 1
RESULT 4
LOCUS AX517113/c
DEFINITION Sequence 3111 from Patent WO02052044.
ACCESSION AX517113
VERSION AX517113.1 GI:23565449
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
  1
  Nakamura,Y., Sekine,A., Iida,A. and Saito,S.
  Detection of genetic polymorphisms
  Patent: WO 02052044-A 3311 04-JUL-2002;
  Riken (JP)
  Location/Qualifiers
  source
    1..40
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
ORIGIN
  Query Match
  Best Local Similarity 3.1%; Score 20.6; DB 6; Length 40;
  Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 133 CAGGAGCAGCTTCAGAAAGTCTGTAAGTTTGTGG 167
Db 36 CAGCAGCAAGCTGGGAGAGTAGCTGCAGATTGTGG 2
RESULT 5
LOCUS AX519642/c
DEFINITION Sequence 5840 from Patent WO02052044.
ACCESSION AX519642
VERSION AX519642.1 GI:23569993
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
  1
  Nakamura,Y., Sekine,A., Iida,A. and Saito,S.
  Detection of genetic polymorphisms
  Patent: WO 02052044-A 5840 04-JUL-2002;
  Riken (JP)
  Location/Qualifiers
  source
    1..40
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
ORIGIN
  Query Match
  Best Local Similarity 3.1%; Score 20.6; DB 6; Length 40;
  Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 133 CAGGAGCAGCTTCAGAAAGTCTGTAAGTTTGTGG 167
Db 36 CAGCAGCAAGCTGGGAGAGTAGCTGCAGATTGTGG 2
RESULT 6
LOCUS E41550/c
DEFINITION Method for yielding male abortive plants.
ACCESSION E41550
```





```
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Exon 11 ('GR2005')"
```

ORIGIN

Query Match  
Best Local Similarity 3.0%; Score 20; DB 6; Length 50;  
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 510 GCTCCTCCAGCCGAGCGTCTCTGATGCGAGCTGAAGC 545  
Db |||||

RESULT 10  
AR075912/c  
LOCUS  
DEFINITION Improved transformation efficiency in phase presentation by  
ACCESSION BD247107 45 bp DNA linear PAT 30-AUG-2000  
VERSION JP 2002522029-A/169  
KEYWORDS synthetic construct  
ORGANISM synthetic construct  
SOURCE artificial sequences  
REFERENCE 1 (bases 1 to 45)  
AUTHORS Sidhu, S.S., Weiss, G.A. and Wells, J.A.  
TITLE Improved transformation efficiency in phase presentation by  
JOURNAL modification of coat proteins  
COMMENT' Patent: JP 2002522029-A 169 23-JUL-2002;  
GENENTECH INC  
OS Artificial Sequence  
PN JP 2002522029-A/169  
PD 23-JUL-2002  
PF 22-JUL-1999 JP 2000562499  
PR 27-JUL-1998 US 60/094291, 08-OCT-1998 US 60/103514 PR  
10-MAY-1999 US 60/133296, 19-MAY-1999 US 60/134870 PI  
SACHDEV S SIDHU, GREGORY A WEISS, JAMES A WELLS PC  
C12N15/09, C07K14/01, C07K14/52, C07K14/715, C07K16/18, C07K19/00, PC  
C12N1/15,  
PC C12N1/19, C12N1/21, C12N5/10, C12N7/00, C12N15/02, C12P21/02, PC  
(C12P21/02, C12P1/92), C12N15/00, C12N15/00, C12N5/00 CC sequence  
source: synthetic  
FT Key Location/Qualifiers  
FH source 1..45  
FT /organism="Artificial Sequence".

FEATURES  
source  
1..45  
Location/Qualifiers  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

ORIGIN

Query Match 2.9%; Score 19.8; DB 6; Length 45;  
Best Local Similarity 69.2%; Pred. No. 1.1e+07;  
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 280 CCAGAGGCGCCCTCCGTTGCCCCCACCAGCGCCGCCCGCCGCGAGCC 318  
Db |||||

RESULT 13  
120781/c  
LOCUS  
DEFINITION Sequence 5 from patent US 5516653.  
ACCESSION 120781 45 bp DNA linear PAT 07-OCT-1996  
VERSION 120781.1 GI:1601136  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 45)  
AUTHORS Bard, J.A., Walker, M.W., Branchek, T. and Weinshank, R.L.  
TITLE DNA encoding a human neuro peptide Y-peptide YY/pancreatic  
polypeptide receptor (Y4) and uses thereof  
JOURNAL Patent: US 5516653-A 5 14-MAY-1996;  
FEATURES Location/Qualifiers  
source 1..45  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 2.9%; Score 19.8; DB 6; Length 45;  
Best Local Similarity 69.2%; Pred. No. 1.1e+07;  
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 131 GACAGGACGCTTCAGAAAGTTGCTGAAGTTTGTGTC 169  
Db |||||

RESULT 11  
AR083198/c  
LOCUS  
DEFINITION Sequence 5 from patent US 5976814.  
ACCESSION AR083198 45 bp DNA linear PAT 01-SEP-2000  
VERSION AR083198.1 GI:10009988  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 45)  
AUTHORS Bard, J.A., Walker, M.W., Branchek, T. and Weinshank, R.L.  
TITLE DNA encoding a human neuro peptide Y-peptide YY/pancreatic  
polypeptide receptor (Y4) and uses thereof  
JOURNAL Patent: US 5976814-A 5 02-NOV-1999;  
FEATURES Location/Qualifiers  
source 1..45  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 2.9%; Score 19.8; DB 6; Length 45;  
Best Local Similarity 69.2%; Pred. No. 1.1e+07;  
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 131 GACAGGACGCTTCAGAAAGTTGCTGAAGTTTGTGTC 169  
Db |||||



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 29, 2004, 11:52:38 ; Search time 267.801 Seconds  
(without alignments)  
9052.319 Million cell updates/sec

Title: US-09-904-568-1\_COPY\_132\_803

Perfect score: 672

Sequence: 1 atgtctgctttggggctgc.....gtgagcgcaactgcaggac 672

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2353733 seqs, 1803733377 residues

Total number of hits satisfying chosen parameters: 1385142

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

1: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/prodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/prodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/prodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/prodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/prodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/prodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/prodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/prodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/prodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/prodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
18: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20.2	3.0	35	13	US-10-109-886-2
2	20	3.0	50	10	Sequence 2, Appli
3	20	3.0	50	10	Sequence 463, App
4	19.8	2.9	42	15	US-10-131-827-677
5	19.6	2.9	50	14	Sequence 677, App
6	19.4	2.9	50	14	US-10-319-315-56
7	19.2	2.9	44	15	US-10-011-931-8
8	19.2	2.9	44	14	Sequence 8, Appli
9	19.2	2.9	44	14	Sequence 309, App
10	19.2	2.9	47	15	Sequence 19, Appl
11	19.2	2.9	50	15	US-10-331-289-19
12	18.8	2.8	50	15	Sequence 1287, Ap
13	18.6	2.8	38	15	US-10-131-827-1362
14	18.6	2.8	38	15	Sequence 1362, Ap
15	18.6	2.8	43	14	US-09-978-917A-23
					Sequence 23, Appl
					Sequence 31, Appl
					Sequence 32, Appl
					Sequence 15, Appl

c	16	18.6	2.7	50	15	US-10-131-827-5574	Sequence 5574, Ap
	17	18.4	2.7	37	9	US-03-996-606-5	Sequence 5, Appl1
	18	18.4	2.7	40	15	US-10-447-839A-78	Sequence 78, Appl1
	19	18.4	2.7	50	9	US-09-989-722-251	Sequence 251, App
	20	18.4	2.7	50	9	US-09-989-723-251	Sequence 251, App
	21	18.4	2.7	50	9	US-09-989-279-251	Sequence 251, App
	22	18.4	2.7	50	9	US-09-989-727-251	Sequence 251, App
	23	18.4	2.7	50	9	US-09-989-731-251	Sequence 251, App
	24	18.4	2.7	50	9	US-09-989-732-251	Sequence 251, App
	25	18.4	2.7	50	9	US-09-991-073-251	Sequence 251, App
	26	18.4	2.7	50	9	US-09-990-442-251	Sequence 251, App
	27	18.4	2.7	50	9	US-09-991-163-251	Sequence 251, App
	28	18.4	2.7	50	9	US-09-993-604-251	Sequence 251, App
	29	18.4	2.7	50	9	US-09-990-456-251	Sequence 251, App
	30	18.4	2.7	50	9	US-09-989-721-251	Sequence 251, App
	31	18.4	2.7	50	9	US-09-992-598-251	Sequence 251, App
	32	18.4	2.7	50	9	US-09-989-293A-251	Sequence 251, App
	33	18.4	2.7	50	9	US-09-989-735-251	Sequence 251, App
	34	18.4	2.7	50	9	US-09-990-444-251	Sequence 251, App
	35	18.4	2.7	50	9	US-09-991-181-251	Sequence 251, App
	36	18.4	2.7	50	9	US-09-989-730-251	Sequence 251, App
	37	18.4	2.7	50	9	US-09-990-436-251	Sequence 251, App
	38	18.4	2.7	50	9	US-09-993-687-251	Sequence 251, App
	39	18.4	2.7	50	10	US-09-989-734-251	Sequence 251, App
	40	18.4	2.7	50	10	US-09-997-653-251	Sequence 251, App
	41	18.4	2.7	50	10	US-09-993-667-251	Sequence 251, App
	42	18.4	2.7	50	10	US-09-997-428-251	Sequence 251, App
	43	18.4	2.7	50	10	US-09-997-666-251	Sequence 251, App
	44	18.4	2.7	50	10	US-09-990-438-251	Sequence 251, App
	45	18.4	2.7	50	10	US-09-990-562-251	Sequence 251, App

#### ALIGNMENTS

#### RESULT 1

US-10-109-886-2  
; Sequence 2, Application US/10109886  
; Publication No. US20020119499A1  
; GENERAL INFORMATION:  
; APPLICANT: TANABE SEIVAKU CO. LTD.  
; APPLICANT: TANIGUCHI, Tomoyasu  
; APPLICANT: MIZUKAMI, Junko  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND  
; FILE REFERENCE: TANIGUCHI=6  
; CURRENT APPLICATION NUMBER: US/10/109,886  
; CURRENT FILING DATE: 2002-04-01  
; PRIOR APPLICATION NUMBER: 09/514,247  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: PCT/JP98/03734  
; PRIOR FILING DATE: 1998-08-24  
; PRIOR APPLICATION NUMBER: JP231084/1997  
; PRIOR FILING DATE: 1997-08-27  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 35  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificially synthesized primer sequence  
US-10-109-886-2

Query Match 3.0%; Score 20.2; DB 13; Length 35;  
Best Local Similarity 75.8%; Pred. No. 2e+04;  
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 86 CCTCTCAGAAAGTACGGCGGCTCAGCTCT 118

1 CCGCTCAGAAATGTTGGCAGTGGCTCAGGACT 33

## RESULT 2

US-09-993-346-463/c  
; Sequence 463, Application US/09993346  
; Publication No. US20030124530A1  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Cynthia A.  
; Cantor, Charles R.  
; Andrews, Beth M.  
; Turin, Lisa M.  
; Fry, Kirk E.

TITLE OF INVENTION: Sequence-Directed DNA Binding  
; Molecules, Compositions and Methods

NUMBER OF SEQUENCES: 664

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genelabs Technologies, Inc.

STREET: 505 Penobscot Drive

CITY: Redwood City

STATE: CA

COUNTRY: USA

ZIP: 94063

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/993,346

FILING DATE: 13-NOV-2000 US20030124530A1-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/354,947

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/171,389

FILING DATE: 20-DEC-1993

APPLICATION NUMBER: US 08/123,936

FILING DATE: 17-SEP-1993

APPLICATION NUMBER: US 07/996,783

FILING DATE: 23-DEC-1992

APPLICATION NUMBER: US 07/723,618

FILING DATE: 27-JUN-1991

APPLICATION NUMBER: US 08/081,070

FILING DATE: 22-JUN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Brady, John F.

REGISTRATION NUMBER: 39,118

REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 324-0880

TELEFAX: (650) 324-0960

INFORMATION FOR SEQ ID NO: 463:

SEQUENCE CHARACTERISTICS:

LENGTH: 50 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: Human insulin-like growth factor II

gene

SEQUENCE DESCRIPTION: SEQ ID NO: 463:

US-09-993-346-463

## Query Match

Best Local Similarity 3.0%; Score 20; DB 10; Length 50;

Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

## QY

259 GGGGGCACACACCGCTCTCCAGCAGGCTCTCCGGCTGCCGCC 302

## Db

49 GGGGGGGCCACCGCCGGCTCTTATATCGGGCCAGCGCGGCC 6

## RESULT 3

US-10-131-827-677

; Sequence 677, Application US/10131827

; Publication No. US20040009479A1

; GENERAL INFORMATION:

; APPLICANT: Wohlgenuth, Jay

; APPLICANT: Fry, Kirk

; APPLICANT: Woodward, Robert

; APPLICANT: Ly, Ngoc

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUN

; FILE REFERENCE: 506612000120

; CURRENT APPLICATION NUMBER: US/10/131,827

; PRIOR FILING DATE: 2002-09-06

; PRIOR APPLICATION NUMBER: US 10/006,290

; PRIOR FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: US 60/296,764

; PRIOR FILING DATE: 2001-06-08

; NUMBER OF SEQ ID NOS: 9090

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 677

; LENGTH: 50

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-131-827-677

Query Match

Best Local Similarity 3.0%; Score 20; DB 15; Length 50;

Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

## QY

172 GCCTGTCATGGAAGACTGCAGAGAAG 199

## Db

16 GCGGGATGGGAAGACAGAGAAG 43

## RESULT 4

US-10-319-315-56

; Sequence 56, Application US/10319315

; Publication No. US20030219774A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: NOVEL HUMAN NEUROTRANSMITTER TRANSPORTER

; FILE REFERENCE: D0205 NP

; CURRENT APPLICATION NUMBER: US/10/319,315

; CURRENT FILING DATE: 2002-12-13

; NUMBER OF SEQ ID NOS: 135

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 56

; LENGTH: 42

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-319-315-56

Query Match

Best Local Similarity 2.9%; Score 19.8; DB 15; Length 42;

Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

## QY

23 CTCCTACTTGACCATCCGCTGACAGTCACAGTGGCC 61

## Db

3 CCCAAGCTTGCCCATGCCGAAGAAGACGAAAGTACC 41

## RESULT 5

US-10-011-931-8/c

; Sequence 8, Application US/10011931

; Publication No. US20030026806A1

; GENERAL INFORMATION:

; APPLICANT: WITTE, ALISON

; APPLICANT: VARNUM, BRIAN C.

; APPLICANT: QIAN, ZUEMING

; APPLICANT: VEZINA, CHRIS

; TITLE OF INVENTION: ANTIBODIES AND OTHER SELECTIVE IL-1 BINDING AGENTS THAT ALLOW BI

; FILE REFERENCE: A-731

; CURRENT APPLICATION NUMBER: US/10/011,931



APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Chumakov, Ilya  
TITLE OF INVENTION: Biallelic markers for use in constructing a high density...

FILE REFERENCE: GENSET.020CP1  
CURRENT FILING DATE: 2003-01-21  
CURRENT APPLICATION NUMBER: US/10/349,143

PRIOR FILING DATE: 1999-10-20  
PRIOR APPLICATION NUMBER: US/09/422,978

PRIOR FILING DATE: 1999-04-21  
PRIOR APPLICATION NUMBER: US 09/298,850

PRIOR FILING DATE: 1998-11-23  
PRIOR APPLICATION NUMBER: US 60/109,732

PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: US 60/082,614

NUMBER OF SEQ ID NOS: 11796  
SEQ ID NO 2676

LENGTH: 47  
TYPE: DNA

ORGANISM: Homo Sapiens  
FEATURE:

NAME/KEY: allele  
LOCATION: 24

OTHER INFORMATION: 99-13864-64 : polymorphic base G or T  
US-10-349-143-2676

Query Match 2.9%; Score 19.2; DB 15; Length 47;  
Best Local Similarity 70.6%; Pred. No. 4.7e+04;

Matches 24; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 272 CCCTGCTCCAGCAGCTCTCCGGCTGCCCTCCCTGC 305

DB 10 CCCTCTCCACAGCTCCCTCCCTCCCTCCCTGC 43

RESULT 10  
US-10-131-827-1287

Sequence 1287, Application US/10131827  
Publication No. US20040009479A1

GENERAL INFORMATION:  
APPLICANT: Wohlgemuth, Jay

APPLICANT: Fry, Kirk  
APPLICANT: Woodward, Robert

APPLICANT: Ly, Ngoc  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE

FILE REFERENCE: 0219us310 - protein C  
CURRENT FILING DATE: 2001-10-17

NUMBER OF SEQ ID NOS: 48  
SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 23  
LENGTH: 50

TYPE: DNA  
ORGANISM: Artificial Sequence

FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Primer

US-09-978-917A-23

Query Match 2.8%; Score 18.8; DB 10; Length 50;  
Best Local Similarity 63.0%; Pred. No. 6.3e+04;

Matches 29; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 465 CCGTGGCGGGTGGAGCTGCGCCATCTCAACGAGCTCAGTCCCG 510

DB 1 CCCTGCGAGGTGGTCTGCTGCACTCAACCAAGAGCTGCGCTGC 46

RESULT 13  
US-10-138-195-31

Sequence 31, Application US/10138195  
Publication No. US20030207340A1

GENERAL INFORMATION:  
APPLICANT: Morrie, D. James

APPLICANT: Morrie, Dorothy J.  
APPLICANT: Chueh, Pin-ju

TITLE OF INVENTION: Sequences Encoding Human Neoplastic Marker  
FILE REFERENCE: 85-99

CURRENT APPLICATION NUMBER: US/10/138,195  
CURRENT FILING DATE: 2002-05-01

RESULT 11  
US-10-131-827-1362/c

Sequence 1362, Application US/10131827

Publication No. US20040009479A1  
GENERAL INFORMATION:  
APPLICANT: Wohlgemuth, Jay  
APPLICANT: Fry, Kirk  
APPLICANT: Woodward, Robert  
APPLICANT: Ly, Ngoc  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE

FILE REFERENCE: 506612000120  
CURRENT FILING DATE: 2002-09-06

PRIOR FILING DATE: 2001-10-22  
PRIOR APPLICATION NUMBER: US 10/006,290

PRIOR FILING DATE: 2001-06-08  
PRIOR APPLICATION NUMBER: US 60/296,764

NUMBER OF SEQ ID NOS: 9090  
SOFTWARE: Patent In version 3.1

SEQ ID NO 1362  
LENGTH: 50

TYPE: DNA  
ORGANISM: Homo sapiens

US-10-131-827-1362

Query Match 2.9%; Score 19.2; DB 15; Length 50;  
Best Local Similarity 67.5%; Pred. No. 4.7e+04;

Matches 27; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 178 CATGGGAAGACTGAGAGAGCTGTGGAGCACTTGGTG 217

DB 42 CAGTGGGAAGAAATGCAAGAAAGCTTTAGAGACATAGATG 3

RESULT 12  
US-09-978-917A-23

Sequence 23, Application US/09978917A  
Publication No. US20030027299A1

GENERAL INFORMATION:  
APPLICANT: Maxygen Aps; Maxygen Holdings

TITLE OF INVENTION: Protein C or activated protein C-like molecules  
FILE REFERENCE: 0219us310 - protein C

CURRENT FILING DATE: 2001-10-17  
NUMBER OF SEQ ID NOS: 48

SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 23

LENGTH: 50  
TYPE: DNA

ORGANISM: Artificial Sequence  
FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Primer  
US-09-978-917A-23

Query Match 2.8%; Score 18.8; DB 10; Length 50;  
Best Local Similarity 63.0%; Pred. No. 6.3e+04;

Matches 29; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 465 CCGTGGCGGGTGGAGCTGCGCCATCTCAACGAGCTCAGTCCCG 510

DB 1 CCCTGCGAGGTGGTCTGCTGCACTCAACCAAGAGCTGCGCTGC 46

RESULT 13  
US-10-138-195-31

Sequence 31, Application US/10138195  
Publication No. US20030207340A1

GENERAL INFORMATION:  
APPLICANT: Morrie, D. James

APPLICANT: Morrie, Dorothy J.  
APPLICANT: Chueh, Pin-ju

TITLE OF INVENTION: Sequences Encoding Human Neoplastic Marker  
FILE REFERENCE: 85-99

CURRENT APPLICATION NUMBER: US/10/138,195  
CURRENT FILING DATE: 2002-05-01

; PRIOR APPLICATION NUMBER: US 60/162,644  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR APPLICATION NUMBER: PCT US00/30190  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 31  
; LENGTH: 38  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: oligonucleotide  
; OTHER INFORMATION: useful as primer, for example.  
US-10-138-195-31

Query Match 2.8%; Score 18.6; DB 15; Length 38;  
Best Local Similarity 72.7%; Pred. No. 6.7e+04;  
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 358 CAGGATCTAATTGGAGATTGGCCAGTTGGCA 390  
Db 1 CAGGAATGACTGGAGTTGTGGCCAGCCTGGAA 33

RESULT 14  
US-10-138-195-32/c  
; Sequence 32, Application US/10138195  
; Publication No. US20030207340A1  
; GENERAL INFORMATION:  
; APPLICANT: Morre, D. James  
; APPLICANT: Morre, Dorothy J.  
; APPLICANT: Chueh, Pin-Ju  
; TITLE OF INVENTION: Sequences Encoding Human Neoplastic Marker  
; FILE REFERENCE: 85-99  
; CURRENT APPLICATION NUMBER: US/10/138,195  
; CURRENT FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: US 60/162,644  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR APPLICATION NUMBER: PCT US00/30190  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 32  
; LENGTH: 38  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: oligonucleotide  
; OTHER INFORMATION: useful as primer, for example.  
US-10-138-195-32

Query Match 2.8%; Score 18.6; DB 15; Length 38;  
Best Local Similarity 72.7%; Pred. No. 6.7e+04;  
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 358 CAGGATCTAATTGGAGATTGGCCAGTTGGCA 390  
Db 38 CAGGAATGACTGGAGTTGTGGCCAGCCTGGAA 6

RESULT 15  
US-10-176-255-15  
; Sequence 15, Application US/10176255  
; Publication No. US20030153004A1  
; GENERAL INFORMATION:  
; APPLICANT: Behan, Dominic P.  
; APPLICANT: Chalmers, Derek T.  
; APPLICANT: Liaw, Chen W.  
; APPLICANT: Russo, Joseph F.  
; APPLICANT: Thomsen, William J.  
; TITLE OF INVENTION: No. US20030153004A1-Endogenous, Constitutively Activated Human Se  
; TITLE OF INVENTION: Small Molecule Modulators Thereof  
; FILE REFERENCE: AREN-0328

; CURRENT APPLICATION NUMBER: US/10/176,255  
; CURRENT FILING DATE: 2002-06-19  
; PRIOR APPLICATION NUMBER: 09/060,188  
; PRIOR FILING DATE: 1998-04-14  
; PRIOR APPLICATION NUMBER: 09/767,013  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 09/292,072  
; PRIOR FILING DATE: 1999-04-14  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 15  
; LENGTH: 43  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic construct  
US-10-176-255-15

Query Match 2.8%; Score 18.6; DB 14; Length 43;  
Best Local Similarity 65.9%; Pred. No. 6.9e+04;  
Matches 27; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
QY 391 TTGGAGTCAACGCCCTCTTCTCGACTCTGTAGCCCAACA 431  
Db 1 TTCAGCAGTCAACCCACTAGTCTATCTCTGTCAACAAA 41

Search completed: February 29, 2004, 14:51:49  
Job time : 270.801 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 29, 2004, 10:05:33 ; Search time 2724.97 Seconds  
(without alignments)  
10688.739 Million cell updates/sec

Title: US-09-904-568-1\_COPY\_132\_803  
Perfect score: 672  
Sequence: 1 atgtctgctttgggggctgc.....gtggcgcaactgcaggac 672

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 1603530

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_em.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_ov.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_in.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rnd.\*
- 36: em\_htg\_man.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	21.8	3.2	49	6	AX404880	AX404880 Sequence
C 2	21.4	3.2	50	6	BD174394	BD174394 Cell usef
C 3	21.2	3.2	47	6	AR284756	AR284756 Sequence
C 4	21	3.1	40	6	E14011	E14011 Probe 7/19
C 5	20.8	3.1	42	6	AX710949	AX710949 Sequence
C 6	20.8	3.1	42	6	BD001090	BD001090 Method an
C 7	20.8	3.1	42	6	BD001519	BD001519 Method an
C 8	20.6	3.1	41	6	AX514286	AX514286 Sequence
C 9	20.6	3.1	41	6	AX520468	AX520468 Sequence
C 10	20.6	3.1	47	6	AR284587	AR284587 Sequence
C 11	20.2	3.0	35	6	E22807	E22807 Method for
C 12	20.2	3.0	35	6	AR203329	AR203329 Sequence
C 13	20	3.0	50	6	AR032851	AR032851 Sequence
C 14	20	3.0	50	6	I29591	I29591 Sequence 46
C 15	20	3.0	50	6	I91265	I91265 Sequence 46
C 16	20	3.0	50	6	AR209515	AR209515 Sequence
C 17	20	3.0	50	8	AJ600712	AJ600712 Arabidops
C 18	19.8	2.9	46	6	AR157914	AR157914 Sequence
C 19	19.4	2.9	50	6	AX199530	AX199530 Sequence
C 20	19.2	2.9	32	6	BD061703	BD061703 Antigenic
C 21	19.2	2.9	44	6	AR149584	AR149584 Sequence
C 22	19.2	2.9	47	6	AR290941	AR290941 Sequence
C 23	19	2.8	27	6	AR089960	AR089960 Sequence
C 24	19	2.8	27	6	AR196995	AR196995 Sequence
C 25	19	2.8	27	6	AR259149	AR259149 Sequence
C 26	18.8	2.8	42	6	AX080016	AX080016 Sequence
C 27	18.8	2.8	43	6	AX466471	AX466471 Sequence
C 28	18.8	2.8	48	10	S77040	S77040 T-cell rece
C 29	18.6	2.8	50	6	BD225137	BD225137 Polymorph
C 30	18.6	2.8	43	6	AR107092	AR107092 Sequence
C 31	18.6	2.8	43	6	AR111855	AR111855 Sequence
C 32	18.6	2.8	43	6	AR118372	AR118372 Sequence
C 33	18.6	2.8	43	6	AR302214	AR302214 Sequence
C 34	18.6	2.8	43	6	AX813374	AX813374 Sequence
C 35	18.6	2.8	43	6	BD132482	BD132482 Binding m
C 36	18.6	2.8	47	6	AX236856	AX236856 Sequence
C 37	18.4	2.7	31	6	AX286510	AX286510 Sequence
C 38	18.4	2.7	37	6	AR012326	AR012326 Sequence
C 39	18.4	2.7	37	6	AR012348	AR012348 Sequence
C 40	18.4	2.7	39	6	BD065047	BD065047 Hematopoi
C 41	18.4	2.7	40	6	E49257	E49257 Omega 3 fat
C 42	18.4	2.7	40	6	AX538550	AX538550 Sequence
C 43	18.4	2.7	47	6	I77232	I77232 Sequence 22
C 44	18.4	2.7	48	6	AX582839	AX582839 Sequence
C 45	18.4	2.7	50	6	AR252548	AR252548 Sequence

ALIGNMENTS

RESULT 1	AX404880/c	AX404880	49 bp	DNA	linear	PAT 14-JUN-2002
LOCUS	AX404880	Sequence 13 from Patent WO0222833.				
DEFINITION	AX404880	AX404880.1	GI:21438112			
ACCESSION	AX404880					
VERSION	AX404880.1					
KEYWORDS		synthetic construct				
SOURCE		synthetic construct				
ORGANISM		artificial sequences.				
REFERENCE		1				
AUTHORS		Pfizenmaier,K., Wuest,T., Moosmayer,D., Grell,M. and Scheurich,P.				
TITLE		Fusion protein from antibody cytokine-cytokine inhibitor				
		(selectokine) for use as target-specific prodrug				
JOURNAL		Patent: WO 0222833-A 13 21-MAR-2002;				

Universitaet Stuttgart (DE) ; Pfizenmaier, Klaus (DE)  
FEATURES  
source  
Location/Qualifiers  
1..49  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Primer 4 fuer die Amplifikation eines humanen  
TNF-Fragments"

ORIGIN  
Query Match 3.2%; Score 21.8; DB 6; Length 49;  
Best Local Similarity 70.7%; Pred. No. 8.1e+05;  
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 74 TGGGTTCCAGCCCTCTCCAGAGTGACGGCGGTGCTCAG 114  
DB 46 TGGGATCATTCCTCCGTCCGGATCGAGGTCGTGATCCG 6

RESULT 2  
LOCUS BD174394 50 bp DNA linear PAT 18-MAR-2003  
DEFINITION Cell useful for evaluation of action of sex hormone receptor.  
ACCESSION BD174394  
VERSION BD174394.1 GI:29120078  
KEYWORDS JP 2002247986-A/27.  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 50)  
AUTHORS Nakao, T., Matsue, K. and Takebayashi, N.  
TITLE Cell useful for evaluation of action of sex hormone receptor  
JOURNAL Patent: JP 2002247986-A 27 03-SEP-2002;  
MITSUI CHEMICALS INC  
COMMENT OS Artificial Sequence  
PN JP 2002247986-A/27  
PD 03-SEP-2002  
PF 23-FEB-2001 JP 2001048876  
PI TOSHIFUMI NAKAO, KAORI MATSUE, NOZOMI TAKEBAYASHI PC  
C12N15/09, C12N5/10, C12Q1/66, C12Q1/66, C12N5/10, C12R1/91, PC  
(C12Q1/02, C12R1/91), (C12Q1/66, C12R1/91), C12N15/00, C12N5/00, PC  
(C12N5/00, C12R1/91)  
CC Description of Artificial Sequence:Primer  
FH Key Location/Qualifiers  
FT source 1..50  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

ORIGIN  
Query Match 3.2%; Score 21.4; DB 6; Length 50;  
Best Local Similarity 66.0%; Pred. No. 1e+06;  
Matches 31; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
QY 136 AGCAGCTTCAGAAAGTTCTTGAAGTGTAGTCGGGCGCTGCTGG 182  
DB 4 AGCTTCGCGACAGCTTTCAGCTGTAAAGCGGGTCTGCTGAGGG 50

RESULT 3  
LOCUS AR284756/c 47 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 808 from patent US 6528260.  
ACCESSION AR284756  
VERSION AR284756.1 GI:29721660  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 47)

Universitaet Stuttgart (DE) ; Blumenfeld, M., Chumakov, I., Bougueleret, L. and Cohen, A.  
FEATURES  
source  
Location/Qualifiers  
1..47  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 3.2%; Score 21.2; DB 6; Length 47;  
Best Local Similarity 65.9%; Pred. No. 1.2e+06;  
Matches 29; Conservative 1; Mismatches 14; Indels 0; Gaps 0;  
QY 96 AGTGAGGCGCGTCTTTCAGCTCTTGAAGGACTTAGACAGGCA 139  
DB 44 AGCGATGCCCTTGGGCGAGAKTTGGAGGGAAGGGGACAGGCA 1

RESULT 4  
LOCUS E14011 40 bp DNA linear PAT 28-JUL-1999  
DEFINITION Probe.  
ACCESSION E14011  
VERSION E14011.1 GI:5708694  
KEYWORDS JP 1997257798-A/1.  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 40)  
AUTHORS Shimada, K. and Namatame, Y.  
TITLE IMMOBILIZATION OF GENE  
JOURNAL Patent: JP 1997257798-A 1 03-OCT-1997;  
SUMITOMO METAL IND LTD  
COMMENT OS None  
PN JP 1997257798-A/1  
PD 03-OCT-1997  
PF 19-MAR-1996 JP 1996062885  
PI SHIMADA KAZUNORI, NAMATAME YASUKO  
PC G01N33/566, C12N15/09, C12Q1/68;  
CC strandedness: Single;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
FH Key Location/Qualifiers  
FT source 1..40  
/organism="Artificial sequences".

FEATURES  
source  
Location/Qualifiers  
1..40  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

ORIGIN  
Query Match 3.1%; Score 21; DB 6; Length 40;  
Best Local Similarity 73.0%; Pred. No. 1.3e+06;  
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 308 GTCTAAAGCCAGATGCTTCCAGGAGAGCTCCAGCA 344  
DB 4 GTCTCAGCTCATTCGCTTTCAGGTAGAGCTCTTGA 40

RESULT 5  
LOCUS AX710949 42 bp RNA linear PAT 11-APR-2003  
DEFINITION Sequence 249 from Patent EP1288296.  
ACCESSION AX710949  
VERSION AX710949.1 GI:29787330  
KEYWORDS Human herpesvirus 5  
SOURCE Human herpesvirus 5  
ORGANISM Human herpesvirus 5

Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
Betaherpesvirinae; Cytomegalovirus.

## REFERENCE

AUTHORS Draper,K.G., Mcswiggen,J.A., Holecek,J.J., Dudycz,L.W.,  
Macejak,D.G. and Mamone,A.J.  
TITLE Method and reagent for inhibiting HBV viral replication  
JOURNAL Patent: EP 1288296-A 249 05-MAR-2003;  
RIBOZYME PHARMACEUTICALS, INC. (US)

## FEATURES

source 1. .42  
/organism="Human herpesvirus 5"  
/mol\_type="unassigned RNA"  
/db\_xref="taxon:10359"

## ORIGIN

Query Match 3.1%; Score 20.8; DB 6; Length 42;  
Best Local Similarity 70.0%; Pred. No. 1.5e+06;  
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 201 TGTGGAGCAACTTGTGCGCCGACCTGTCTCAGAGAG 240

Db 2 TGTGGATGACCTACGGGCCATCGCGGAGGAGTCAGATGAG 41

## RESULT 6

BD001090 42 bp RNA linear PAT 31-JAN-2002  
LOCUS Method and reagent for inhibiting viral replication.

ACCESSION BD001090

VERSION BD001090.1 GI:18625649

KEYWORDS JP 2000342285-A/250.

SOURCE synthetic construct

ORGANISM artificial sequences.

REFERENCE 1 (bases 1 to 42)

## AUTHORS

Draper,K.G., Dadykztz,L.W., Macswigen,J.A., Maysejak,D.G.,  
Holesek,J.J. and Mamone,A.J.

TITLE Method and reagent for inhibiting viral replication

JOURNAL Patent: JP 2000342285-A 250 12-DEC-2000;

COMMENT RIBOZYME PHARMACEUTICALS INC

OS Artificial Sequence

PN JP 2000342285-A/250

PF 12-DEC-2000

PR 11-MAY-1992 US 07/882689,14-MAY-1992 US 07/882712 PR

14-MAY-1992 US 07/882713,14-MAY-1992 US 07/882714 PR

14-MAY-1992 US 07/882823,14-MAY-1992 US 07/882824 PR

14-MAY-1992 US 07/882886,14-MAY-1992 US 07/882888 PR

14-MAY-1992 US 07/882922,14-MAY-1992 US 07/882921 PR

14-MAY-1992 US 07/882889,14-MAY-1992 US 07/883823 PR

14-MAY-1992 US 07/883849,14-MAY-1992 US 07/884073 PR

14-MAY-1992 US 07/884074,14-MAY-1992 US 07/884333 PR

14-MAY-1992 US 07/884422,14-MAY-1992 US 07/884431 PR

14-MAY-1992 US 07/884436,14-MAY-1992 US 07/884521 PR

31-JUL-1992 US 07/923738,26-AUG-1992 US 07/935854 PR

26-AUG-1992 US 07/936086,18-SEP-1992 US 07/948359 PR

15-OCT-1992 US 07/963322,07-DEC-1992 US 07/987129 PR

07-DEC-1992 US 07/987130,07-DEC-1992 US 07/987133 PI

KENNETH G DRAPER,LEC W DADYKTZ,JAMES A MACSWIGEN, PI DENNIS G

MAYSEJAK,

PI JAMES J HOLESEK,ANTHONY J MAMONE

PC C12N15/09,C12N5/10,C12N7/00,C12N9/22/(C12N5/10,C12R1:91), PC

C12N15/00,

PC C12N5/00,(C12N5/00,C12R1:91)

CC C12N5/00,

PH Key Location/Qualifiers

FT source 1. .42

FEATURES Location/Qualifiers

source 1. .42

/organism="synthetic construct"

/mol\_type="genomic RNA"

/db\_xref="taxon:32630"

## ORIGIN

Query Match 3.1%; Score 20.8; DB 6; Length 42;  
Best Local Similarity 70.0%; Pred. No. 1.5e+06;  
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 201 TGTGGAGCAACTTGTGCGCCGACCTGTCTCAGAGAG 240

Db 2 TGTGGATGACCTACGGGCCATCGCGGAGGAGTCAGATGAG 41

## RESULT 7

BD001519 42 bp RNA linear PAT 31-JAN-2002  
LOCUS Method and reagent for inhibiting viral replication.

DEFINITION BD001519

ACCESSION BD001519

VERSION BD001519.1 GI:18626078

KEYWORDS JP 2000342286-A/250.

SOURCE synthetic construct

ORGANISM synthetic construct

artificial sequences.

REFERENCE 1 (bases 1 to 42)

AUTHORS Draper,K.G., Dadykztz,L.W., Macswigen,J.A., Maysejak,D.G.,

Holesek,J.J. and Mamone,A.J.

TITLE Method and reagent for inhibiting viral replication

JOURNAL Patent: JP 2000342286-A 250 12-DEC-2000;

COMMENT RIBOZYME PHARMACEUTICALS INC

OS Artificial Sequence

PN JP 2000342286-A/250

PD 12-DEC-2000

PF 01-MAY-2000 JP 2000132651

PR 11-MAY-1992 US 07/882689,14-MAY-1992 US 07/882712 PR

14-MAY-1992 US 07/882713,14-MAY-1992 US 07/882714 PR

14-MAY-1992 US 07/882823,14-MAY-1992 US 07/882824 PR

14-MAY-1992 US 07/882886,14-MAY-1992 US 07/882888 PR

14-MAY-1992 US 07/882889,14-MAY-1992 US 07/882921 PR

14-MAY-1992 US 07/883823,14-MAY-1992 US 07/883823 PR

14-MAY-1992 US 07/883849,14-MAY-1992 US 07/884073 PR

14-MAY-1992 US 07/884074,14-MAY-1992 US 07/884333 PR

14-MAY-1992 US 07/884422,14-MAY-1992 US 07/884431 PR

14-MAY-1992 US 07/884436,14-MAY-1992 US 07/884521 PR

31-JUL-1992 US 07/923738,26-AUG-1992 US 07/935854 PR

26-AUG-1992 US 07/936086,18-SEP-1992 US 07/948359 PR

15-OCT-1992 US 07/963322,07-DEC-1992 US 07/987129 PR

07-DEC-1992 US 07/987130,07-DEC-1992 US 07/987133 PI

KENNETH G DRAPER,LEC W DADYKTZ,JAMES A MACSWIGEN, PI DENNIS G

MAYSEJAK,

PI JAMES J HOLESEK,ANTHONY J MAMONE

PC C12N15/09,C12N5/10,C12N7/00/(A61K38/43,A61K39/125,A61K39/13,

PC A61K39/135,

PC A61K39/145,A61K39/21,A61K39/23,A61K39/245,A61K39/29,A61K48/00,

PC A61P1/16,

PC A61P31/14,A61P31/16,A61P31/18,A61P31/22,A61P35/02,C12O1/68, PC

(C12N15/09,C12R1:93),C12N15/00,C12N5/00,A61K37/48,(C12N15/00, PC

C12R1:93)

CC

PH Key Location/Qualifiers

FT source 1. .42

FEATURES Location/Qualifiers

source 1. .42

/organism="synthetic construct"

/mol\_type="genomic RNA"

/db\_xref="taxon:32630"

## ORIGIN

Query Match 3.1%; Score 20.8; DB 6; Length 42;  
Best Local Similarity 70.0%; Pred. No. 1.5e+06;  
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 201 TGTGGAGCAACTTGTGCGCCGACCTGTCTCAGAGAG 240

Db 2 TGTGGATGACCTACGGGCCATCGCGGAGGAGTCAGATGAG 41

RESULT 8  
AX514286/c  
LOCUS AX514286 41 bp DNA linear PAT 05-OCT-2002  
DEFINITION Sequence 484 from Patent WO02052044.  
ACCESSION AX514286  
VERSION AX514286.1 GI:23560673  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Nakamura, Y., Sekine, A., Iida, A. and Saito, S.  
TITLE Detection of genetic polymorphisms  
JOURNAL Patent: WO 02052044-A 484 04-JUL-2002;  
Riken (JP)  
FEATURES  
source  
1..41  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 3.1%; Score 20.6; DB 6; Length 41;  
Best Local Similarity 70.3%; Pred. No. 1.7e+06;  
Matches 26; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
QY 20 CAGCTCCATCTGACCATCCGCTGACAGTCACAG 56  
|||||  
Db 38 CAGCCCTGTAACGACATCTGCTGAGGCTCTCGG 2  
|||||  
RESULT 9  
AX520468/c  
LOCUS AX520468 41 bp DNA linear PAT 05-OCT-2002  
DEFINITION Sequence 6666 from Patent WO02052044.  
ACCESSION AX520468  
VERSION AX520468.1 GI:23571065  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Nakamura, Y., Sekine, A., Iida, A. and Saito, S.  
TITLE Detection of genetic polymorphisms  
JOURNAL Patent: WO 02052044-A 6666 04-JUL-2002;  
Riken (JP)  
FEATURES  
source  
1..41  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 3.1%; Score 20.6; DB 6; Length 41;  
Best Local Similarity 70.3%; Pred. No. 1.7e+06;  
Matches 26; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
QY 20 CAGCTCCATCTGACCATCCGCTGACAGTCACAG 56  
|||||  
Db 38 CAGCCCTGTAACGACATCTGCTGAGGCTCTCGG 2  
|||||  
RESULT 10  
AR284587  
LOCUS AR284587 47 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 639 from patent US 6528260.  
ACCESSION AR284587  
VERSION AR284587.1 GI:29721491  
KEYWORDS

SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 47)  
AUTHORS Blumenfeld, M., Chumakov, I., Bougueleret, L. and Cohen, A.  
TITLE Biallelic markers related to genes involved in drug metabolism  
JOURNAL Patent: US 6528260-A 639 04-MAR-2003;  
FEATURES  
source  
1..47  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 3.1%; Score 20.6; DB 6; Length 47;  
Best Local Similarity 64.4%; Pred. No. 1.7e+06;  
Matches 29; Conservative 1; Mismatches 15; Indels 0; Gaps 0;  
QY 596 AGCTGGGTACAGTGTAGCTTGGTCCTTAAGGAGATGCCAGAAC 640  
|||||  
Db 3 AGTAGGGGACAGTGAAGCATKGGGGGGCCAGGAGTCCAGAGC 47  
|||||  
RESULT 11  
E22807  
LOCUS E22807 35 bp DNA linear PAT 18-JUN-2001  
DEFINITION Method for screening PPAR agonist and antagonist.  
ACCESSION E22807  
VERSION E22807.1 GI:13024121  
KEYWORDS JP 1999056369-A/2.  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 35)  
AUTHORS Junko, M. and Tomoyasu, T.  
TITLE Method for screening PPAR agonist and antagonist  
JOURNAL Patent: JP 1999056369-A 2 02-MAR-1999;  
TANABE SEIYAKU CO LTD  
COMMENT OS Unidentified  
PN JP 1999056369-A/2  
PD 02-MAR-1999  
PF 27-AUG-1997 JP 1997231084  
PR  
PI JUNKO MINAKAMI, TOMOYASU TANIGUCHI  
PC C12N15/09; C12N1/19, C12Q1/68 // (C12N1/19, C12R1/645), C12N15/00 CC  
Strandedness: Single;  
CC Topology: Linear;  
FH Key Location/Qualifiers  
FT source 1..35  
FT /organism="Unidentified".  
FEATURES  
source  
1..35  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Query Match 3.0%; Score 20.2; DB 6; Length 35;  
Best Local Similarity 75.8%; Pred. No. 2.2e+06;  
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 86 CCTCTCCAGAGTGACGGCGCTGGCTCAGCTCT 118  
|||||  
Db 1 CCGCTCGAGAAATGTTGGCAGTGGCTCAGGACT 33  
|||||  
RESULT 12  
AR203329  
LOCUS AR203329 35 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 2 from patent US 6365361.  
ACCESSION AR203329  
VERSION AR203329.1 GI:21499690  
KEYWORDS  
SOURCE Unknown.

```
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 35)
AUTHORS Taniguchi, T. and Mizukami, J.
TITLE Method for identifying or screening agonist and antagonist to PPAR
JOURNAL Patent: US 6365361-A 2 02-APR-2002;
FEATURES
    source
        1..35
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
    Query Match          3.0%; Score 20.2; DB 6; Length 35;
    Best Local Similarity 75.8%; Pred. No. 2.2e+06;
    Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 86 CCTCTCCAGAGTGACGCGCGCTGCTCAGCTCT 118
Db 1 CCGCTCGAGAAATGTTGGCAGTGGCTCAGGACT 33
RESULT 13
LOCUS AR032851/c
DEFINITION Sequence 463 from patent US 5869241.
ACCESSION AR032851
VERSION AR032851.1 GI:5948456
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 50)
AUTHORS Edwards, C.A., Cantor, C.R., Andrews, B.M., Turin, L.M. and Fry, K.E.
TITLE Method of determining DNA sequence preference of a DNA-binding molecule
JOURNAL Patent: US 5869241-A 463 09-FEB-1999;
FEATURES
    source
        1..50
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
    Query Match          3.0%; Score 20; DB 6; Length 50;
    Best Local Similarity 65.9%; Pred. No. 2.4e+06;
    Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 259 GCGGGCACACACACCTGCTCCAGCAGGCTCTCCGGTGCCTCC 302
Db 49 GCGGGCGCCACCGCGGCTCTTATAGTCGCCGACCGCGGCC 6
RESULT 14
LOCUS 129591/c
DEFINITION Sequence 463 from patent US 5578444.
ACCESSION 129591
VERSION 129591.1 GI:1820382
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 50)
AUTHORS Edwards, C.A., Cantor, C.R., Andrews, B.M., Turin, L.M. and Fry, K.E.
TITLE Sequence-directed DNA-binding molecules compositions and methods
JOURNAL Patent: US 5578444-A 463 26-NOV-1996;
FEATURES
    source
        1..50
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
    Query Match          3.0%; Score 20; DB 6; Length 50;
    Best Local Similarity 65.9%; Pred. No. 2.4e+06;
    Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 259 GCGGGCACACACACCTGCTCCAGCAGGCTCTCCGGTGCCTCC 302
Db 49 GCGGGCGCCACCGCGGCTCTTATAGTCGCCGACCGCGGCC 6
MATCHES 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 259 GCGGGCACACACACCTGCTCCAGCAGGCTCTCCGGTGCCTCC 302
Db 49 GCGGGCGCCACCGCGGCTCTTATAGTCGCCGACCGCGGCC 6
RESULT 15
LOCUS 191265/c
DEFINITION Sequence 463 from patent US 5726014.
ACCESSION 191265
VERSION 191265.1 GI:3935735
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 50)
AUTHORS Edwards, C.A., Cantor, C.R., Andrews, B.M. and Turin, L.M.
TITLE Screening assay for the detection of DNA-binding molecules
JOURNAL Patent: US 5726014-A 463 10-MAR-1998;
FEATURES
    source
        1..50
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
    Query Match          3.0%; Score 20; DB 6; Length 50;
    Best Local Similarity 65.9%; Pred. No. 2.4e+06;
    Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 259 GCGGGCACACACACCTGCTCCAGCAGGCTCTCCGGTGCCTCC 302
Db 49 GCGGGCGCCACCGCGGCTCTTATAGTCGCCGACCGCGGCC 6
MATCHES 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 259 GCGGGCACACACACCTGCTCCAGCAGGCTCTCCGGTGCCTCC 302
Db 49 GCGGGCGCCACCGCGGCTCTTATAGTCGCCGACCGCGGCC 6
Search completed: February 29, 2004, 13:32:14
Job time : 2731.97 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 29, 2004, 10:34:43 ; Search time 2031.49 Seconds  
(without alignments)  
9878.168 Million cell updates/sec

Title: US-09-904-568-1\_COPY\_132\_803

Perfect score: 672

Sequence: 1 atgtctgcttggggctgc.....gtgagcgcaactgcaggac 672

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 138346

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	22	3.3	50	9	AU106355
C 2	20.6	3.1	50	9	AU104945
C 3	20.4	3.0	40	9	AA933650
C 4	20.4	3.0	49	28	AZ971265
					AZ971265 2M0244N12

C	5	20.4	3.0	50	9	AU106354
C	6	19.8	2.9	50	9	AU104331
C	7	19.6	2.9	47	29	AL757838
C	8	19.4	2.9	41	12	BI669410
C	9	19.4	2.9	47	12	BI544737
C	10	19.4	2.9	50	9	AU106356
C	11	19.4	2.9	50	12	BG314884
C	12	19.2	2.9	50	9	AU104702
C	13	19	2.8	42	28	AQ025866
C	14	19	2.8	50	28	AZ479659
C	15	18.8	2.8	50	9	AU106353
C	16	18.8	2.8	50	9	AU107055
C	17	18.8	2.8	50	9	AU107064
C	18	18.6	2.8	50	9	AU103136
C	19	18.4	2.7	40	28	AZ666390
C	20	18.4	2.7	42	12	BJ063892
C	21	18.4	2.7	44	12	BM400948
C	22	18.4	2.7	45	12	BM397151
C	23	18.4	2.7	49	9	AA280723
C	24	18.4	2.7	50	29	CG720601
C	25	18.2	2.7	40	12	BI825771
C	26	18.2	2.7	43	9	AA116340
C	27	18.2	2.7	46	9	AA052407
C	28	18.2	2.7	48	29	CG802067
C	29	18.2	2.7	50	9	AU103070
C	30	18.2	2.7	50	9	AU104084
C	31	18	2.7	43	12	BJ001433
C	32	18	2.7	45	28	BZ585215
C	33	18	2.7	47	9	AV833049
C	34	18	2.7	47	28	AZ778260
C	35	18	2.7	47	28	BZ596680
C	36	18	2.7	50	9	AU102993
C	37	18	2.7	50	9	AU103858
C	38	18	2.7	50	9	AU103860
C	39	18	2.7	50	9	AU103898
C	40	18	2.7	50	9	AU104250
C	41	18	2.7	50	9	AU105404
C	42	18	2.7	50	9	AU107645
C	43	18	2.7	50	9	AU107646
C	44	17.8	2.6	38	28	AQ025526
C	45	17.8	2.6	41	12	BG390320

## ALIGNMENTS

RESULT 1	AU106355	50 bp	mrna	linear	EST 30-AUG-2001
LOCUS	AU106355	Sugano Homo sapiens cdna library	Homo sapiens	cdna clone	
DEFINITION	KAT02155	mrna sequence.			
ACCESSION	AU106355				
VERSION	AU106355.1	GI:13555876			
KEYWORDS	BST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 50)				
AUTHORS	Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.				
TITLE	Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites				
JOURNAL	EMBO Rep. 2 (5), 388-393 (2001)				
MEDLINE	21270072				
PUBMED	11375929				
COMMENT	Contact: Yutaka Suzuki Department of Virology Institute of Medical Science, University of Tokyo 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan Email: ysuzuki@ims.u-tokyo.ac.jp Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and				

Sugano, S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

#### FEATURES

Location/Qualifiers  
1..50  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="KAT02155"  
/clone\_lib="Sugano Homo sapiens cDNA library"

#### ORIGIN

Query Match 3.3%; Score 22; DB 9; Length 50;  
Best Local Similarity 73.7%; Pred. No. 3e+05;  
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 264 CACACACCCGCTCCAGCAGCTCTCCGGTGGCCCC 301  
|||||  
DB 50 CCACGACCCAGCAGCTCCGGCTCTCCAGCGCCCC 13  
|||||

#### RESULT 2

AU104945/c  
LOCUS AU104945 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone  
DEFINITION HRC05721, mRNA sequence.  
ACCESSION AU104945  
VERSION AU104945.1 GI:13554466  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

#### REFERENCE

1 (bases 1 to 50)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J.,  
Rata, H., Oca, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K.,  
Sakaki, Y., Nakamura, Y., Suyama, A. and Sugano, S.  
Diverse transcriptional initiation revealed by fine, large-scale  
mapping of mRNA start sites  
EMBO Rep. 2 (5), 388-393 (2001)

#### JOURNAL

#### MEDLINE

#### PUBMED

#### COMMENT

Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: ysuzuki@ims.u-tokyo.ac.jp  
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and  
Sugano, S. Construction and characterization of a full  
length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),  
149-156 (1997).

#### FEATURES

Location/Qualifiers  
1..50  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HRC05721"  
/clone\_lib="Sugano Homo sapiens cDNA library"

#### ORIGIN

Query Match 3.1%; Score 20.6; DB 9; Length 50;  
Best Local Similarity 74.3%; Pred. No. 6.6e+05;  
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 270 CACCTGCTCCAGCAGGCTCTCCGGTCCCGCTG 304  
|||||  
DB 45 CACTGCTCCAGCGGCTCCAGCGTCCCGCTGCTG 11  
|||||

#### RESULT 3

AA933650/c  
LOCUS AA933650 40 bp mRNA linear EST 27-APR-1998  
DEFINITION OM56C06.s1 NCI\_CGAP\_GC4 Homo sapiens cDNA clone IMAGE:1551178 3'

similar to SW:DPD2\_HUMAN P49005 DNA POLYMERASE DELTA SMALL SUBUNIT  
; mRNA sequence.

ACCESSION AA933650  
VERSION AA933650.1 GI:3089918  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

#### JOURNAL

#### COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Seq primer: -40m13 fwd. BT from Amersham

High quality sequence stop: 1.

#### FEATURES

#### source

Location/Qualifiers  
1..40  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1551178"  
/tissue\_type="pooled germ cell tumors"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP GC4"  
/notes="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; 1st strand cDNA was prepared from 3 pooled  
germ cell tumors, and was then primed with a Not I -  
oligo(dT) primer. Double-stranded cDNA was ligated to Eco  
RI adaptors (Pharmacia), digested with Not I and cloned  
into the Not I and Eco RI sites of the modified pT73  
vector. Library is normalized. Library was constructed by  
Bento Soares and M. Fatima Bonaldo."

#### ORIGIN

Query Match 3.0%; Score 20.4; DB 9; Length 40;  
Best Local Similarity 71.1%; Pred. No. 6.5e+05;  
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 ATGTCTGCTTTGGGGCTGCAGCTCCATCTTCACCA 38  
|||||  
DB 39 ATGTCTACTTTTGTGGCACACCCCGCTTGTCTCCA 2  
|||||

#### RESULT 4

#### LOCUS

#### DEFINITION

#### ACCESSION

#### VERSION

#### KEYWORDS

#### SOURCE

#### ORGANISM

#### REFERENCE

#### AUTHORS

1 (bases 1 to 49)  
Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,  
Rilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von  
Niederhausern, A. and Wright, D., Weiss, R.

AZ971265 49 bp DNA linear GSS 27-APR-2001  
2M0244N12R Mouse 10kb plasmid UUGC2M library Mus musculus genomic  
clone UUGC2M0244N12 R, genomic survey sequence.

AZ971265 GI:13842492

GSS.

**TITLE**  
Mouse whole genome scaffolding with paired end reads from 10kb  
**JOURNAL**  
Unpublished (2000)  
**COMMENT**  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLCT, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: rdunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0244 row: N column: 12  
Seq primer: CACACAGGAAACAGCATGTGCC  
Class: Plasmid ends  
High quality sequence stop: 49.  
Location/Qualifiers  
1..49  
    /organism="Mus musculus"  
    /mol\_type="genomic DNA"  
    /strain="C57BL/6J"  
    /db\_xref="taxon:10090"  
    /cclone="UUGC2M0244N12"  
    /sex="Female"  
    /lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"  
    /cloned\_lib="Mouse 10kb plasmid UUGC2M library"  
    /note=vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (female) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
    (http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adapted DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pWD42 [gil|4732114|gb|AF129072.1], a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adapted mouse DNA was annealed to  
adapted vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

**FEATURES**  
                source  
1..49  
    /organism="Mus musculus"  
    /mol\_type="genomic DNA"  
    /strain="C57BL/6J"  
    /db\_xref="taxon:10090"  
    /cclone="UUGC2M0244N12"  
    /sex="Female"  
    /lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"  
    /cloned\_lib="Mouse 10kb plasmid UUGC2M library"  
    /note=vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (female) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
    (http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adapted DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pWD42 [gil|4732114|gb|AF129072.1], a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adapted mouse DNA was annealed to  
adapted vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

**ORIGIN**

Query Match         3.0%     Score 20.4; DB 28; Length 49;  
Best Local Similarity     71.1%; Pred.No. 7.3e+05;  
Matches   27; Conservative   0; Mismatches   11; Indels    0; Gaps    0;

**QY**   252 CCTGCTGGCGGCACACACACCCTGTCCAGCAGGCTC 289  
         ||| |||||  
**Db**    42 CCAGAAGCGCGCACCATATGCTTCTCCAGCAGTGTC 5  
         ||| |||||

**RESULT 5**  
AUI016354/c  
LOCUS  
DEFINITION  
COL00324, mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 50)  
Suzuki,Y., Iaira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,  
Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,  
Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.  
Diverse transcriptional initiation revealed by fine, large-scale



LOCUS BI669410 41 bp mRNA linear EST 12-SEP-2001  
DEFINITION 60329525F1 NIH\_MGC\_96 Homo sapiens cDNA clone IMAGE:5314750 5',  
mRNA sequence.  
ACCESSION BI669410  
VERSION BI669410.1 GI:15583643  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 41)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLM11797 row: m column: 23  
High quality sequence stop: 41.  
Location/Qualifiers  
1. 41  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5314750"  
/tissue\_type="hypothalamus"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_96"  
/notes="Organ: Brain; Vector: pBluescriptR (modified  
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI  
(gtcgag); Oligo-dT primed using primer  
5'-TTTTTTTTTTTTTNN-3', size-selected for average  
insert size 2.3 kb and normalized to 10^5. This is a  
primary library enriched for full-length clones and  
constructed using the Cap-trapper method (Carninci, in  
preparation). Library constructed by M. Brownstein  
(NIH/NHGRI, National Institutes of Health). Note: this is  
a NIH\_MGC Library."

ORIGIN  
Query Match 2.9%; Score 19.4; DB 12; Length 41;  
Best Local Similarity 79.3%; Pred. No. 1.2e+06;  
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 283 CAGGCTCTCCGCTGCCCGCTGCTAGTCT 311  
DB 41 CAAGATGTCGCGCGCGCGCGCTGCTGCT 13

RESULT 9  
BI544737/c  
LOCUS BI544737 47 bp mRNA linear EST 05-SEP-2001  
DEFINITION 603242001F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:5284488 5',  
mRNA sequence.  
ACCESSION BI544737  
VERSION BI544737.1 GI:15432049  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 47)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.

LOCUS BI669410 47 bp DNA linear GSS 18-JUN-2002  
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-152H09-013139,  
genomic survey sequence.  
ACCESSION AL757838  
VERSION AL757838.1 GI:21496186  
KEYWORDS GSS.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsi.  
REFERENCE 1  
AUTHORS Strizhov, N., Li, Y., Rosso, M., Viehoveer, P., Dekker, K., Saedler, H.  
and Weisshaar, B.  
TITLE A pipeline for automated high-throughput generation of FSTs  
(flanking sequence tags) from Arabidopsis thaliana T-DNA  
transformed lines  
JOURNAL Unpublished  
REFERENCE 2  
AUTHORS Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weisshaar, B.  
TITLE A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)  
for flanking sequence tag based reverse genetics  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 47)  
AUTHORS Strizhov, N., Li, Y., Rosso, M. and Weisshaar, B.  
TITLE Direct Submission  
JOURNAL Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer  
Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany  
COMMENT This sequence is recovered from the left border of the T-DNA. It  
indicates an insertion close to or within gene At4gl8700. The  
sequences are generated at the MPI for Plant Breeding Research in  
the context of the GABI-Kat project. GABI-Kat is part of the German  
Plant Genomics program designated 'GABI'. Information on line  
availability can be found at:  
<http://www.mpiz-koeln.mpg.de/GABI-Kat/>.  
Location/Qualifiers  
1. 47  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia 0"  
/db\_xref="taxon:3702"  
/clone="GK-152H09-013139"  
/clone\_lib="Arabidopsis thaliana T-DNA insertion lines"  
/note="PCR was performed on DNA from Arabidopsis thaliana  
plants (T1) which were transformed with the T-DNA from  
vector pAC161. The lines contain one or more T-DNA from  
insertions. The DNA fragment(s) resulting from the PCR  
were directly sequenced to determine the genomic sequence  
flanking the insertion. Sequences displaying significant  
similarity to the A. thaliana nuclear genome sequence were  
processed for submission. T-DNA derived sequences were  
removed"

ORIGIN  
Query Match 2.9%; Score 19.6; DB 29; Length 47;  
Best Local Similarity 73.5%; Pred. No. 1.1e+06;  
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 571 GAGTCCCATACCCAAATTCAGGAGCTCGCGT 604  
DB 39 GAGGATCCCATAGGTTTGTTCGGAGCTCCCGT 6

RESULT 8  
BI669410/c

Email: csapba-remail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L2A411719 row: a column: 01  
 High quality sequence stop: 47.  
 Location/Qualifiers  
 1. 47  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:528488"  
 /tissue\_type="hippocampus"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_95"  
 /note="Organ: brain; Vector: pBluescriptR (modified  
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI  
 (Gcgag); Oligo-dT primed using primer  
 5'-TTTTTTTTTTTTTNN-3', size-selected for average  
 insert size 2.5 kb and normalized to ROT 5. This is a  
 primary library enriched for full-length clones and  
 constructed using the Cap-trapper method (Carninci, in  
 preparation). Library constructed by M. Brownstein  
 (NIH/NHGRI, National Institutes of Health). Note: this  
 is a NIH\_MGC Library."

## ORIGIN

Query Match 2.9%; Score 19.4; DB 12; Length 47;  
 Best Local Similarity 79.3%; Pred. No. 1.3e+06;  
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 283 CAGGCTCTCGGTCGCCCTGCTAGTCT 311

Db 47 CAAGATGTCGCCGCCCTGCTAGTCT 19

RESULT 10  
 AU106356/c  
 LOCUS  
 DEFINITION AU106356 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone  
 KAT03974, mRNA sequence.

ACCESSION AU106356  
 VERSION AU106356  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)  
 AUTHORS Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J.,  
 Kata, H., Oka, T., Isozaki, T., Tanaka, T., Morishita, S., Okubo, K.,  
 Sakaki, Y., Nakamura, Y., Suyama, A. and Sugano, S.

TITLE Diverse transcriptional initiation revealed by fine, large-scale  
 mapping of mRNA start sites

JOURNAL EMBO Rep. 2 (5), 388-393 (2001)

MEDLINE 21270072

PubMed 11375929

COMMENT Contact: Yutaka Suzuki  
 Department of Virology  
 Institute of Medical Science, University of Tokyo  
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
 Email: ysuzuki@ims.u-tokyo.ac.jp

Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and  
 Sugano, S. Construction and characterization of a full  
 length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),  
 149-156 (1997).

## FEATURES

source Location/Qualifiers  
 1. 50

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="KAT03974"  
 /clone\_lib="Sugano Homo sapiens cDNA library"

## ORIGIN

Query Match 2.9%; Score 19.4; DB 9; Length 50;  
 Best Local Similarity 70.3%; Pred. No. 1.3e+06;  
 Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 258 GCGCGGCACACACACCTGCTCCAGCAGGCTCTCCGG 294

Db 37 GGCACCCCCACACACAGGACTGCCGGCTCTCCAG 1

## RESULT 11

BG314884

LOCUS

DEFINITION BG314884 50 bp mRNA linear EST 14-MAR-2002

ACCESSION BG314884

VERSION BG314884.1 GI:18997731

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)

AUTHORS Andersson, T., Borang, S., Larsson, M., Thelin, A.,  
 Ekstrand-Hammarstrom, B., Wirta, V., Wenborg, A., Lundberg, J. and  
 Odeberg, J.

TITLE Identification of candidate genes in atherosclerosis - Virtual chip  
 analysis in RDA based transcript profiling of monocyte/macrophage  
 response to oxidised LDL

JOURNAL Unpublished (2001)

COMMENT Contact: Andersson Tove

Department of Biotechnology

KTH

Teknikringen 34, plan 6, 100 44 Stockholm, Sweden

Tel: +46 8 790 71 29

Fax: +46 8 245452

Email: tove@biochem.kth.se

POLYA-No.

source Location/Qualifiers

1. 50

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone\_lib="THP1"

/cell\_line="THP1"

/note="Vector: pRIT28; Site 1: BamHI; Site 2: BamHI;

Shotgun cloning of RDA difference products. Macrophage and  
 foamcell libraries were submitted to successive rounds of  
 subtractive hybridisations generating populations of gene  
 fragments that are differentially expressed in macrophage  
 to foam cell formation."

## ORIGIN

Query Match 2.9%; Score 19.4; DB 12; Length 50;  
 Best Local Similarity 68.4%; Pred. No. 1.3e+06;  
 Matches 26; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 170 GGGCCCTGCAATGGGAAGACTGCAGAGAGCTGTGGAG 207

Db 7 GAGCCGTGCTTGGTAGGCACGGGAGTGAAGCTGGGAG 44

## RESULT 12

AU104702

LOCUS

DEFINITION AU104702 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone

AU104702

LOCUS

DEFINITION HRC06060, mRNA sequence.

ACCESSION AU104702  
 VERSION AU104702.1 GI:13554223  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 REFERENCE 1 (bases 1 to 50)  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Hata, H., Oca, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki, Y., Nakamura, Y., Suyama, A. and Sugano, S.  
 TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites  
 JOURNAL EMBO Rep. 2 (5), 388-393 (2001)  
 MEDLINE 21270072  
 PUBMED 11375929  
 COMMENT Contact: Yutaka Suzuki  
 Department of Virology  
 Institute of Medical Science, University of Tokyo  
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
 Email: ysuzuki@ims.u-tokyo.ac.jp  
 Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano, S. Construction and Characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).  
 FEATURES source  
 1..50  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="HRC06060"  
 /clone\_lib="Sugano Homo sapiens cDNA library"  
 ORIGIN  
 Query Match 2.9%; Score 19.2; DB 9; Length 50;  
 Best Local Similarity 67.5%; Pred. No. 1.5e+06;  
 Matches 27; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
 QY 244 CTGGCGGCTCTGCTGGCGGCGACACACACCTGCTCCAGC 283  
 Db 11 CTGGCGGCGGCGGCGGCGGCTCCAGCAGCGCGCTGC 50  
 RESULT 13  
 AQ025866/c  
 LOCUS  
 DEFINITION Drosophila melanogaster (fruit fly)  
 ACCESSION AQ025866  
 VERSION AQ025866.1 GI:3266218  
 KEYWORDS GSS.  
 SOURCE Drosophila melanogaster  
 ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE 1 (bases 1 to 42)  
 AUTHORS Spradling, A.C., Stern, D., Beaton, A., Rehm, E.J., Laverty, T., Mozdzen, N., Misra, S. and Rubin, G.M.  
 TITLE The BDGP gene disruption project: Single P element insertions mutating 30% of Drosophila autosomal genes  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Gerald Rubin  
 Berkeley Drosophila Genome Project  
 University of California, Berkeley  
 LSA Building, Berkeley, CA 94720-3200, USA  
 Fax: 5106433947  
 Email: gerry@fruitfly.berkeley.edu  
 Sequence recovery method was inverse PCR.  
 Sequence orientation is forward strand relative to 5' end of P element  
 The P element insertion position is base 035 in the 42 bases. This insertion position refers to the first base of the 8 base target recognition sequence.  
 Class: transposon-tagged.  
 FEATURES source  
 1..42  
 Location/Qualifiers  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clone\_lib="Drosophila melanogaster P lethal line"  
 /notes="Inverse PCR was performed on Drosophila melanogaster strains each of which contains a single P transposable element insertion that is thought to cause either lethality or sterility. The resultant fragment for each strain was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at [http://fruitfly.berkeley.edu/p\\_disrupt/inverse\\_pcr.html](http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html)."  
 ORIGIN  
 Query Match 2.8%; Score 19; DB 28; Length 42;  
 Best Local Similarity 69.4%; Pred. No. 1.5e+06;  
 Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
 QY 21 AGCTCCATCTTGCACCATCCGCTGCACAGTCACAG 56  
 Db 36 AGTACTATATGGCTCCATNCCGATACCGCCACAG 1  
 RESULT 14  
 AZ479659/c  
 LOCUS  
 DEFINITION 1M0300B03R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0300B03 R, genomic survey sequence.  
 ACCESSION AZ479659  
 VERSION AZ479659.1 GI:10639488  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 50)  
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D. Weiss, R.  
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0300 row: B column: 03  
 Seq primer: CACACAGAAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 50.  
 FEATURES source  
 1..50  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC1M0300B03"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, TI-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"

